ENERGY DENSITY METHOD AND ITS APPLICATIONS TO DEFECT ENERGIES

BY

MIN YU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Physics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Professor David M. Ceperley, Chair
Professor Tai-Chang Chiang
Professor Richard M. Martin, Advisor
Assistant Professor Dallas R. Trinkle, Advisor

Abstract

We propose a method to study various defects using an energy density formulation in density functional theory. Unlike traditional total energy methods that find a single energy as the difference between two calculations, this approach can provide the formation energies for more than one point defect, surface or interface in one calculation. The energy density method also provides a picture of the distribution of the energy among the surrounding atoms; although the value assigned each atom is not unique, it gives trends. One purpose of this work is to investigate the extent to which the distribution of the energy around the defect can be used to understand its properties. Numerically, we implement the energy density method in the framework of the Vienna ab initio simulation package for the projector augmented wave and pseudopotential methods.

The assignment of energy to an atom uses procedures related to Bader's work on "Atoms in Molecules". Although our work is different, one aspect is the same: the calculation of volume around each atom where the kinetic energy is unique. We extend the idea of Bader volume to find a charge neutral volume for the unique classical Coulomb energy. An important part of our development is a new method to construct volumes in a way that is more accurate and efficient than previous grid-based methods that require very fine grids. Our weight method works with the density on a grid and assigns volume fractions of the cell of each grid point to different atoms. For a wide range of applications, the weight method provides much precise calculations of the atomic energies with a $O(N)$ computing time.

The generality of the method is demonstrated by applications to surface and point defects of semiconductors, metals, and metal oxides. The energy density perturbation around Si

monovacancy reveals the strong energy concentration on {110} planar zigzag chains. The O interstitial in the hexagonal-close-packed Ti crystal demonstrates a Friedel oscillation in both charge density and energy density. In particular we study the interface of Au/TiO_2, which is a well-recognized chemical catalyst at low temperature. By comparing the work of adhesion of different interfaces at the zero temperature, we predict a stable interfacial geometry of Au(111) on the top of a new proposed TiO model with net work of adhesion of $45\,\mathrm{meV/\AA^2}$ and interfacial distance 2.45 Å. Atomic energy variation during forming interfaces demonstrates that the attraction of top Au interfacial layer leads to a stable structure.

To my family

Acknowledgments

My heartfelt thanks to my thesis advisors, Professor Richard M. Martin and Professor Dallas R. Trinkle for their numerous suggestions, patient guidance, and support at all times.

Special gratitude goes to my thesis committee members, Professor David M. Ceperley and Professor Tai-Chang Chiang for their invaluable comments from my prelim to final defense.

I would like to thank my collaborators, Dr. Shankar Sivaramakrishnan and Professor Jian-Min Zuo for their experimental study of the Au/TiO_2 interface.

I should also thank all the professors in the physics department for teaching me all kinds of physics and thank all the other members in electronic structure group and Trinkle's group for useful conversations. I have greatly benefited from discussions with them.

Finally, I sincerely thank my grandparents, my parents, my husband and numerous friends who endured this long process with me, always offering support and love.

I was supported for two years by Schlumberger Foundation Scholarship and one year by the University of Illinois Fellowship. The last year of this research was supported by the National Science Foundation (NSF) under grant number DMR-1006077 and through the Materials Computation Center (MCC) at UIUC, NSF DMR-0325939. Computational resources were provided by the National Center for Supercomputing Applications (NCSA) and the Texas Advanced Computing Center (TACC), and by the MRL Basie cluster and CSE Turing cluster at the University of Illinois.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

a.u.	atomic units
DFT	density functional theory
LDA	the local density approximation in DFT
GGA	the generalized gradient approximation in DFT
NCPP	norm-conserving pseudopotential
USPP	ultrasoft pseudopotential
PAW	the projector augmented wave method
$\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$	many-body electron wavefunction
$\psi(\mathbf{r})$	Kohn-Sham wavefunction
$\hat{n}$	normal direction to surface
$\tilde{\rho}(\mathbf{r})$	pseudo electronic charge density
$\hat{\rho}(\mathbf{r})$	compensation charge density
$\rho^e(\mathbf{r})$	sum of pseudo electronic and compensation charge density
Ω_{cell}	volume of the simulation cell
$\hat{H}$	Hamiltonian
T	kinetic energy in Kohn-Sham theory
E_{XC}	exchange-correlation energy in Kohn-Sham theory
E_{CC}	classical Coulomb energy in Kohn-Sham theory
$t^{(a)}(\mathbf{r})$	real-space antisymmetric form of kinetic energy density
$t^{(s)}(\mathbf{r})$	real-space symmetric form of kinetic energy density

$e_{\mathrm{CC}}^{(a)}(\mathbf{r})$	antisymmetric form of classical Coulomb energy density
$e_{\mathrm{CC}}^{\mathrm{Maxwell}}(\mathbf{r})$	symmetric form of classical Coulomb energy density
$e_{\mathrm{XC}}(\mathbf{r})$	exchange-correlation energy density
$Y_{\ell m}$	spherical harmonic
$R_{\ell m}$	real spherical harmonic

Chapter 1

Introduction

An idealized perfect crystal has a regular atomic structure that extends to infinity in all directions. However, real crystals have surfaces and an array of possible defects that play crucial roles in determining the properties of materials that are important from both scientific and technological point of view. Point defects include vacancies where atoms are missing from the occupied site, and interstitial atoms occupy sites unoccupied by atoms. Line defects extend through the crystal along a one-dimensional line or a one-dimensional curve. Two primary types are edge dislocation and screw dislocation. An edge dislocation can be considered the end of an inserted extra half-plane of atoms; while a screw dislocation is most easily visualized as a spiral in crystal, much like a parking garage. Planar defects occur whenever the crystalline structure of materials is discontinuous crossing a plane, including grain boundary, stacking faults, surfaces, and interfaces between different materials. The difference of atomic size, and number of valence electrons, relaxation of the positions of atoms, and many other effects perturb both structural and electronic properties of materials.

The surface energy of a given facet of a crystal is meaningful for predicting the equilibrium crystal shape and preferred crystal growth directions, and should depend only upon the properties of the surface. The formation energy of a point defect is important for understanding the phase stability, and should depend only upon the properties in the vicinity of the defect. The work of adhesion of an interface is helpful for predicting the equilibrium interface structure, and should depend only upon properties in the interface region during forming interface. In this thesis, one purpose of the work is to decompose the energy of a material into contributions from different spatial regions, so that one can obtain the formation

energies for more than one point defect, surface or interface in one calculation.

1.1 Interfaces of Au nanoparticles on TiO_2(110)

Bulk metallic Au is chemically inert and catalytically inactive as a consequence of combination of valence d orbitals and diffused valence s and p orbitals. Recently, Au nanoparticles have been found to be catalytically active when supported on metal oxides such as TiO_2, SiO_2, Fe_2O_3, Co_3O_4, NiO, Al_2O_3, MgO, etc. [1, 2, 3, 4, 5, 6] For example, Au nanoparticles supported on a TiO_2(110) surface demonstrate catalytic activity to promote the reaction between CO and O_2 to form CO_2 at $T < 40\,\mathrm{K}$ with Au 3.5 nm nanoparticles maximizing activity. [3] The catalytic activity is remarkably sensitive to the support material, Au particle size and Au-support interaction; in addition, the reaction mechanism of CO oxidation over Au/TiO_2 system remains under debate. [3, 7, 8, 9]

High-resolution transmission electron microscopy (HRTEM) [10, 11] and high-angle annular dark field scanning transmission electron microscopy (HAADF-STEM) [11, 12] have characterized the atomic structure of nanocrystal interface. However, the atomic structure of Au/TiO_2 interface is difficult to determine in HRTEM image simulations due to several issues, such as the thickness of nanoparticles and metal oxide substrates are not determined, and the positions of atoms in the direction parallel to the electron beam are not determined. The HAADF mode leads the STEM image intensity highly depending on the atomic number Z. The atomic columns in projection can be directly visualized. Au atoms are clearly visible due to large atomic number, $Z_{\mathrm{Au}} = 79$, compared to $Z_{\mathrm{Ti}} = 22$ and $Z_{\mathrm{O}} = 8$. Therefore, O atoms alone are barely observable. Cosandey *et al.* observed two epitaxial orientations: Au(111) and Au(112) on rutile TiO_2(110) in HRTEM image. [10] Akita *et al.* obtained preferred orientation of Au(111) on rutile TiO_2(110) in HAADF-STEM image. [12] Sivaramakrishnan *et al.* 's HRTEM experiment observed two different epitaxies of Au nanoparticles on TiO_2(110) surfaces: Au(111) and Au(100) as shown in Figure 1.1.

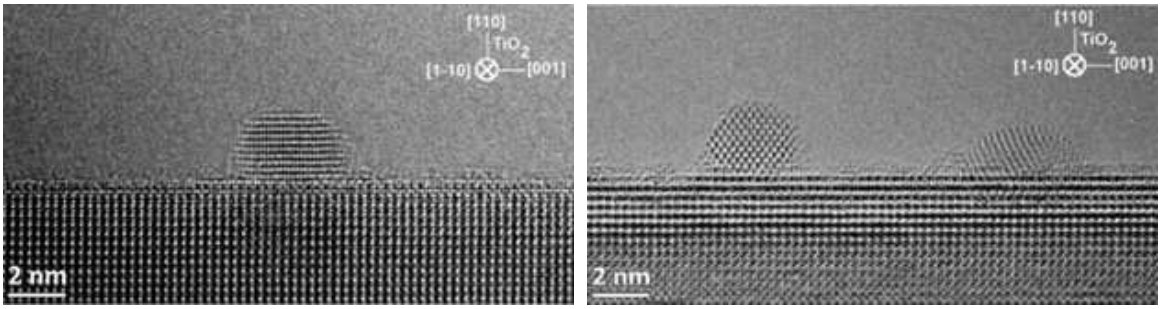

Figure 1.1: HRTEM images [11] of two types epitaxial Au nanoparticles on $TiO_2(110)$ surfaces. Left is $Au(111)//TiO_2(110)$, and right is $Au(100)//TiO_2(110)$.

Au(111) epitaxy is more frequently observed than Au(100). The HAADF-STEM experiment examined a specially reconstructed interface for epitaxial Au(111) sitting on reconstructed $TiO_2(110)$ surface as shown in Figure 1.2. Red and cyan dots indicate the Ti-O columns, and cyan dots indicate Ti-only columns. Two extra O atoms for every Ti atom in Ti-O columns lead to stronger HAADF signals than Ti-only columns. Plot (b) gives the intensity line profile of the HAADF-STEM image in plot (a) inside (blue) and away from (red) the Au nanoparticle. The intensity profile arises across the nanoparticle and drops at the interface (blue curve), while intensity profile off the nanoparticle decreases gradually towards the interface (red curve). The difference of these two curves is shown in plot (b) inset, where the first peak from substrate indicates interfacial layer. This interfacial layer, indicated by the arrow in plot (a), is not a complete Au layer due to the nanopaticle shape, and has stronger signal than TiO_2 alone column. Therefore, it should be a spatially reconstructed interface. The distance between the first fully occupied Au(111) layer and the mixed interfacial layer was measured at 2.35 ± 0.16 Å, which is similar to the distance of bulk Au(111) planes. The distance between the mixed interfacial layer and the next TiO_2 layer was measured at 3.2 ± 0.4 Å, which is similar to the distance of bulk $TiO_2(110)$ planes.

A supported nanocluster with anisotropic surface energies would form a faceted equilibrium shape determined by the minimization of free energy. Figure 1.3 shows STEM image of a Au nanocluster on $TiO_2(110)$ substrate and a schematic representation of a faceted

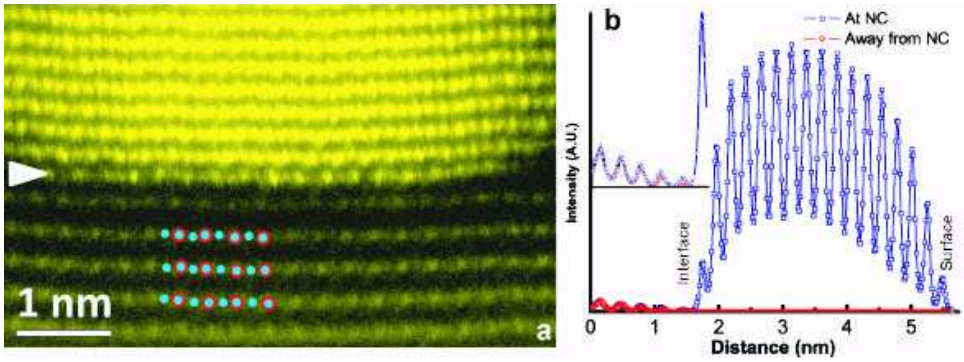

Figure 1.2: HAADF-STEM image and intensity profile [11] of a Au(111) nanoparticle on the TiO_2(110) surface. Red and cyan dots indicate the Ti-O columns, and cyan dots indicate Ti-only columns. Two extra O atoms for every Ti atom in Ti-O columns lead to stronger HAADF signals than Ti-only columns. The arrow points to the Au/TiO_2 interface. Plot (b) shows the intensity line profile of the HAADF-STEM image in plot (a) inside (blue) and away from (red) the Au nanoparticle (averaged size $\sim 2.5\,$nm).

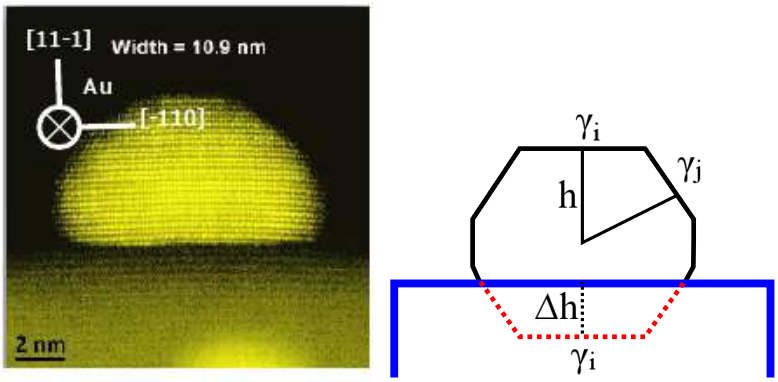

Figure 1.3: STEM image of a Au nanocluster on TiO_2(110) [11] and schematic representation of an equilibrium shape of nanocluster given by the Wulff reconstruction. The Au nanocluster is bounded at the top surface by (111) plane. Part of the nanocluster is truncated at the interface by an amount of Δh.

nanocluster on a substrate with an amount of Δh sunk into the substrate. h is the distance from the center of nanocluster to the top facet. The equilibrium shape of a nanocluster depends on the nanocluster surface energy and work of adhesion of forming interface [13]:

$$\frac{\Delta h}{h} = \frac{E_{\mathrm{adh}}}{\gamma_{\mathrm{Au}}}, \tag{1.1}$$

with work of adhesion defined as $E_{\text{adh}} = \gamma_{\text{Au}} + \gamma_{\text{TiO}_2} - \gamma_{\text{Au/TiO}_2}$. Based on the experimental measurement of geometry parameter $\Delta h/h = 0.65 \pm 0.15$ from nanocluster width ranging from 4 nm to 11 nm and DFT calculation of Au(111) surface energy $\gamma_{\text{Au(111)}} = 43\,\text{meV/Å}^2$ (see details in Section 6.2.2), the work of adhesion of forming Au(111)//TiO_2(110) interface is evaluated as $28 \pm 7\,\text{meV/Å}^2$. To understand the preferential orientation relationship and atomic structure for the Au/TiO_2 interface, it will be effective to combine experimental observation with DFT calculations.

1.2 Background on energy density

In this thesis, we develop an energy density method in density functional theory (DFT) [14] to calculate the *energy distribution* around a defect. The method is closely related to the work of Chetty and Martin [15, 16]; however, here we calculate integrals over the energy density in ways that allow us to go beyond what can be done with those methods. *It must be emphasized from the beginning that there are issues with any method that uses an energy density. The consequences of our work should be clearly separated:*

- *Some energies are rigorously defined with no ambiguity.* This includes all the energies reported here for formation energies of defects and interfaces. Our analysis also shows other well-defined energies. In these cases the value of the method is to provide ways to calculate such quantities with many advantages over traditional total energy methods. For example, in total energy methods, a formation energy is determined as the difference between two calculations, one with and one without the defect. In energy density method, both the ideal bulk energy and the defect energy (or even the energies of more than one defect) can be found in one calculation.

- *The distribution of the energy density in space is not unique, but some aspects are expected to be useful in qualitative and quantitative ways.* One purpose of this thesis is to investigate the extent to which the energy density can be useful in understanding

the properties of defects in addition to their total energy of formation. For example the range of the region around a defect should be given by the approach of the energy to the bulk density. We give several cases where this is useful, and we also investigate further the extent to which our choice of assignment of energies to atoms is useful.

The present work considers the energy density as divided into a kinetic energy density, the total Coulomb energy density including all interactions among electrons and ions; the exchange-correlation energy density in density functional theory which is already defined at each point in space in terms of the density; and other short-ranged terms localized in the core regions around individual atoms. The kinetic and Coulomb energy at any point is not unique; however, the kinetic density is unique except for terms proportional to the Laplacian of electron density similar to a gauge transformation. Nevertheless, gauge independent energies can be obtained as integrals over any volume bounded by a zero-flux (zero perpendicular gradient) surface of the electron density. This aspect of the work is the same defined by Bader [17]; however, we emphasize that other terms are different from those proposed by Bader.

1.3 Atomic energy in Bader's 'Atoms in Molecules' theory

Bader's 'Atoms in Molecules' theory [17] defines atomic subsystems in terms of a property of electron density gradient vector $\nabla\rho$ and then the atomic volume, multipole moments, kinetic and potential energies can be obtained via volume integration. Atoms are separated by interatomic surfaces which sometime can be very complicated and are defined as zero-flux surfaces $\partial\Omega$ of the electron density gradients, $\nabla\rho(\mathbf{r}) \cdot \hat{n}(\mathbf{r}) = 0$. Here, $\rho(\mathbf{r})$ is the electron density; $\hat{n}(\mathbf{r})$ is the unit vector perpendicular to the dividing surface at any surface point $\mathbf{r} \in \partial\Omega$.

The total energy of an atom is a sum of the kinetic energy and the potential energy. The kinetic energy density has two different expressions

$$\begin{aligned} T^{(a)}(\mathbf{r}) &= -\frac{N}{4}\int ... \int [\Psi^*(\mathbf{r},\mathbf{r}_2,...,\mathbf{r}_N)\nabla^2_{\mathbf{r}}\Psi(\mathbf{r},\mathbf{r}_2,...,\mathbf{r}_N) \\ &\quad + \Psi(\mathbf{r},\mathbf{r}_2,...,\mathbf{r}_N)\nabla^2_{\mathbf{r}}\Psi^*(\mathbf{r},\mathbf{r}_2,...,\mathbf{r}_N)]d\mathbf{r}_2...d\mathbf{r}_N\,, \\ T^{(s)}(\mathbf{r}) &= \frac{N}{2}\int ... \int |\nabla_{\mathbf{r}}\Psi(\mathbf{r},\mathbf{r}_2,...,\mathbf{r}_N)|^2 d\mathbf{r}_2...d\mathbf{r}_N\,, \end{aligned} \tag{1.2}$$

which differ by $-\frac{1}{4}\nabla^2\rho(\mathbf{r})$. Ψ is the N-electron wavefunction. The divergence theorem states that the flux of a continuously differentiable vector field $\mathbf{F}$ out of a closed surface $\partial\Omega$ equals to the three dimensional integral of the divergence of the vector field over the volume Ω enclosed by the surface $\partial\Omega$

$$\int\int\int_{\Omega}(\nabla\cdot\mathbf{F})\,dV = \int\int_{\partial\Omega}\mathbf{F}\cdot\hat{n}\,dS\,. \tag{1.3}$$

Therefore, integrated within the atomic volume Ω bounded by zero-flux surfaces of the gradient of electron density, the kinetic energy of an atom is then well defined, $T(\Omega) = T^{(a)}(\Omega) = T^{(s)}(\Omega)$.

In his work, Bader used a local form of virial theorem to relate the local electron kinetic energy $T^{(a)}(\mathbf{r})$ (Eqn. (1.2)) and potential energy density $V(\mathbf{r})$ to the Laplacian of the electronic charge density at any point in space,

$$\frac{1}{4}\nabla^2\rho(\mathbf{r}) = 2\,T^{(a)}(\mathbf{r}) + V(\mathbf{r})\,. \tag{1.4}$$

The electronic energy density can be defined as

$$\begin{aligned} E_e(\mathbf{r}) &= T^{(a)}(\mathbf{r}) + V(\mathbf{r}) \\ &= -\,T^{(s)}(\mathbf{r})\,. \end{aligned} \tag{1.5}$$

He then found the energy assigned to a given volume, Ω bounded by a surface of zero-flux in $\nabla\rho$, solely by the kinetic energy,

$$E_e(\Omega) = -T(\Omega) = \frac{1}{2}V(\Omega)\,. \tag{1.6}$$

This results in charged units (so-called "Bader charges") have long rang interactions. Thus it is arbitrary to assign the energy to one region when it is in fact shared by interacting regions. One can't use Eqn. (1.2) directly in a Kohn-Sham density functional theory [14] as the N-electron wavefunction is unknown; or the exchange-correlation terms are missing if one uses the one-electron wavefunction. A recent paper [18], published during the course of our work, has reported a way to include such terms in a form of the virial theorem; they found different values from those given by the original method. In addition, this approach still has the issue of the way the energies are assigned to regions.

In the present work the Coulomb energy also has a gauge dependence corresponding the expressions in terms of the potential and charge or in terms of the electric field squared. However, like the kinetic energy, a well-defined atomic Coulomb energy can be obtained by integrating in the volume bounded by zero-flux surface of electrostatic field, which is different from the Bader volume. The exchange-correlation energy can be expressed in various functionals, all are given as a function of electron density and its gradients. The short range terms in the core regions are treated differently in different methods; we have derived the appropriate forms for pseudopotential and all-electron PAW methods, which lead to terms assigned to each atom and not explicitly given as densities.

1.4 Organization

The work presented in this thesis is the development and application of a method to decompose the energy of a material into contributions from different spatial regions, so that a single

calculation can provide much more information than the total energy. The generality of the method is demonstrated by application to various defects in crystals, and an extensive study is carried out to provide understanding of the stability and properties of Au/TiO_2 interfaces.

Chapter 2 gives a brief summary of the theoretical concepts of the density functional theory. We review electronic density functional theory, which was first introduced by Hohenberg and Kohn in 1964 [19], and practically implemented by Kohn and Sham in 1965 [14]. We then outline the local density approximation (LDA) [20] and the generalized gradient approximation (GGA) [21, 22] for exchange-correlation energy. The norm-conserving pseudopotential (NCPPs) [23], ultrasoft pseudopotential (USPPs) [24] approximations, and the projector augmented wave method (PAW) [25] are discussed at the end.

Chapter 3 defines the energy density formulae for the NCPPs, the USPPs, and the PAW method. The four terms in total energy density formulation is discussed one by one. We first discuss the kinetic energy densities in general. Follow by the review of Coulomb energy density, and the derivation of our new total Coulomb energy density formalism. Next the general exchange-correlation energy density is explained briefly. The short-ranged terms in the core regions are treated differently in different methods; we give the appropriate forms for pseudopotential and all-electron PAW methods. The chapter's final table encapsulates the total energy density formalism for different methods.

Chapter 4 develops a new weight method that constructs the zero-flux surface of electron density gradients and of potential gradients for unique kinetic energy integration and unique classical Coulomb energy integration, respectively. We describe the volume flowing algorithm of this weight method, and then perform several tests on gauge dependence term integration and Bader charge of model charge, a semiconductor and an ionic compound in terms of grid sizes. We also compare our results with other existing grid-based algorithm in orthogonal lattice and non-orthogonal lattice. The algorithm is robust, efficient with $O(N)$ computing time, and more accurate than other existing grid-based algorithm.

Chapter 5 presents the applications of this method to various defects including: the

atomic displacements in GaAs; the GaAs(110) surface; the Si monovacancy; the Si self-interstitial; the O interstitial in octahedral site in HCP Ti. Each defect energy is calculated from volume integration in the vicinity region of the defect. We compare the results with total energy calculation and experimental data. Furthermore, the energy density method provides a clear picture that energy density pertub around the defect as well as charge density.

Chapter 6 studies the stability and properties of interfaces of Au and TiO_2 by DFT and energy density calculations. We are the first to work on these several different Au/TiO_2 interfaces including: Au(111)//TiO_2(110) and Au(100)//TiO_2(110), with and without bridging O, Au(111) on added-row 1×2 TiO_2(110) reconstruction, and Au(111) on a new proposed TiO 1×2 reconstruction. The analysis of work of adhesion and atomic energy variation during forming interface explains a stable structure Au(111) on the top of TiO reconstruction is relevant to the attraction of top Au interfacial layer.

Atomic units (see Appendix (A)), $\hbar = m_e = e = 4\pi\varepsilon_0 = 1$ are used throughout in equations, while I report computational results in eV and Å.

Chapter 2

Density functional theory

2.1 Introduction

Density functional theory (DFT) has been a primary method in studying the electronic structure in condensed matter since the 1970s. The main objective of DFT is to replace the many-body electron wavefunctions with the electronic charge density as basic variable. The theory was originally developed by two papers: Hohenberg and Kohn (HK) in 1964 [19]; Kohn and Sham (KS) in 1965 [14]. Hohenberg and Kohn introduced a formalism in 1964 that the ground state energy of a many electron system is a unique functional of the electronic charge density $\rho(\mathbf{r})$. Within the Born-Oppenheimer approximation, the many-electron non-relativistic Hamiltonian in atomic units can be written as

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_i V_{\text{ext}}(\mathbf{r}_i) + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}\,. \tag{2.1}$$

The first HK theorem states that the ground state electronic charge density $\rho(\mathbf{r})$ determines the external potential V_{ext} up to a constant, it thus determines the Hamiltonian up to a constant shift of energy, and all other properties of the system. The second HK theorem proves that one can define the total energy of a system as a function of charge density $E[\rho(\mathbf{r})]$. For any external potential V_{ext}, the minimization of total energy occurs at the ground state charge density. Allowing the determination of ground state charge density and ground state energy, HK theorems however have not provided a practical way to determine these quantities. Kohn and Sham [14] proposed an approach in 1965: *to replace the original*

many-body problem by an auxiliary independent-particle problem. It provides a key step of practical calculation and makes DFT a widely used method.

Section 2.2 describes the Kohn-Sham approach. Section 2.3 addresses the fermion feature of interacting electrons by use of exchange-correlation functional. Section 2.4 reviews the pseudopotential and the projector augmented wave (PAW) methods in DFT.

2.2 The Kohn-Sham equations

Kohn and Sham [14] replace the many-body problem of interacting electrons in an external potential with a problem of non-interacting electrons moving in an effective potential. The effective potential includes the external potential and an exchange-correlation functional for all the many-electron effects. The Hohenberg-Kohn energy functional and charge density then depend on one-electron wave functions and their fractional occupations. The standard form of a many-body system total energy is given by

$$\begin{aligned} E_{\mathrm{KS}} = &-\frac{1}{2}\sum_{n\mathbf{k}} f_{n\mathbf{k}} \int d\mathbf{r}\, \psi^*_{n\mathbf{k}}(\mathbf{r})\nabla^2\psi_{n\mathbf{k}}(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r} \int d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \\ &+ \int d\mathbf{r}\rho(\mathbf{r})V_{\mathrm{ext}}(\mathbf{r}) + E_{\mathrm{XC}} + E_{\mathrm{II}}\,, \end{aligned} \tag{2.2}$$

where $\psi_{n\mathbf{k}}(\mathbf{r})$ and $f_{n\mathbf{k}}$ are the Kohn-Sham wavefunction and the electron occupation number for the n^{th} band, for wavevectors $\mathbf{k}$ within the first Brillouin zone. The first term is a non-interacting electron kinetic energy term. The electron density $\rho(\mathbf{r})$ is defined as

$$\rho(\mathbf{r}) = \sum_{n\mathbf{k}} f_{n\mathbf{k}}|\psi_{n\mathbf{k}}|^2\,, \tag{2.3}$$

which is a fundamental measurable, and well-defined quantity. The second term in Eqn. (2.2) is the classical Coulomb interaction energy of electron densities. $V_{\mathrm{ext}}(\mathbf{r})$ is the potential contains the ionic and other external potentials. The exchange-correlation energy E_{XC} (discussed in Section 2.3) describes the many-body effects of electrons. E_{II} is the ion-ion interaction

energy.

The total energy minimization with respect to the wave functions $\psi(\mathbf{r})$ gives the Kohn-Sham Schrödinger-like equations:

$$\left(-\frac{1}{2}\nabla^2 + V_{\rm H}(\mathbf{r}) + V_{\rm ext}(\mathbf{r}) + V_{\rm XC}(\mathbf{r})\right)\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r})\,, \tag{2.4}$$

where ϵ_i are the Kohn-Sham eigenvalues. $V_{\rm H}$ is the Hartree potential of electrons given by

$$V_{\rm H} = \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,, \tag{2.5}$$

which leads to a divergent Hartree energy for a system under periodic boundary condiction. It is thus important to group Hartree potential and ionic potential together for a neutral system. $V_{\rm XC}(\mathbf{r})$ is the exchange-correlation potential (discussed in Section 2.3). The Kohn-Sham equations need to be solved self-consistently.

2.3 Exchange-correlation energy

The fermion nature of many body interacting electrons are expressed in exchange-correlation term. The many-body wavefunction of electrons is antisymmetric under exchange of any two electrons. A spatial separation between two same spin electrons exists due to Pauli exclusion principle and thus reduces the Coulomb energy of the electron system, so called the exchange energy. The Coulomb energy can be even lower when electrons with opposite spin are also spatially separated with the cost of increasing of kinetic energy. The correlation energy can be defined as the difference of exact energy of many body system and the Hartree-Fock approximation.

It is not possible to determine the exchange-correlation energy analytically. In density functional theory, the exchange-correlation energy is usually treated as a function of the local electron density and its gradients, determined by the choice of exchange-correlation

functional. The most widely used approximation is local density approximation (LDA) which assumes $\varepsilon_{\rm XC}(\mathbf{r})$, the exchange-correlation energy per electron at point $\mathbf{r}$ in the electron gas depends only upon the local electron density $\rho(\mathbf{r})$ in the neighbor of point $\mathbf{r}$, and is equal to the exchange-correlation energy per electron in a homogeneous electron gas that has the same electron density $\rho(\mathbf{r})$.

$$E_{\rm XC}^{\rm LDA}(\mathbf{r}) = \int d\mathbf{r}\, \rho(\mathbf{r})\varepsilon_{\rm XC}(\mathbf{r})\,, \tag{2.6}$$

with $\varepsilon_{\rm XC}(\mathbf{r}) = \varepsilon_{\rm XC}^{\rm hom}[\rho(\mathbf{r})]$. Within LDA, the exact expression for the exchange energy of a homogeneous gas is known as $\varepsilon_{\rm X}^{\rm hom} = -3k_{\rm F}/4\pi$, with the wavevector $k_{\rm F} = (3\pi^2\rho)^{1/3}$. But correlation energy is an analytical fit to the numerical quantum Monte Carlo calculation of Ceperley and Alder [26]. Different parameterized formula exist for exchange-correlation energy of a homogeneous electron gas, and lead to very similar total energy results. Two widely used functionals are the PZ scheme of Perdew and Zunger [20] and the VMN scheme of Vosko, Wilk, and Nusiar [27]. In this work, all LDA calculations utilize the PZ parameterized formula. Although LDA ignores the correction of inhomogeneity of electron density, it works remarkably well for a broad range of systems. The success of LDA can be attributed to the exchange-correlation hole obeying a sum rule, $\int \rho_{\rm XC}(\mathbf{r}, \mathbf{r}' - \mathbf{r})d\mathbf{r}' = -1$. [28, 29]

Many attempts to improve on the LDA functional by use of the second-order gradient expansions fail as violating the sum rule. Among those who preserve the sum rule and improve accuracy, the most successful has been the generalized-gradient approximation(GGA), notably PW91 of Perdew and Wang [21] and PBE of Perdew, Burke, and Ernzerhof [22]. The exchange-correlation energy in a GGA scheme is still local but takes into account the gradient of the electron density,

$$\begin{aligned} E_{\rm XC}^{\rm GGA}(\mathbf{r}) &= \int d\mathbf{r}\, \rho(\mathbf{r})\varepsilon_{\rm XC}[\rho, \nabla\rho] \\ &= \int d\mathbf{r}\, \rho(\mathbf{r})\varepsilon_{\rm X}^{\rm hom} F_{\rm XC}[\rho, s]\,, \end{aligned} \tag{2.7}$$

where $s = |\nabla\rho|/(2k_{\rm F}\rho)$ is a dimensionless density gradient in dimensionless parameterized function $F_{\rm XC}[\rho, s]$. Unlike the LDA scheme, many different approximations of $F_{\rm XC}[\rho, s]$ have been proposed for calculations on specific physical systems. They lead to similar improvement on total energies, atomization energies, energy barriers and structural energy difference for most physical systems with small density gradient contribution as functionals have nearly identical shape for small s. The PW91 functional starts from the second-order gradient expansion for the hole $\rho_{\rm XC}$ surrounding an electron and cuts off its long-range contributions to restore exclusion principle. The PW91 functional has been used widely in many physical systems, but it is over-parameterized and has spurious wiggles in exchange-correlation potential $V_{\rm XC}(\mathbf{r}) = \frac{\delta E_{\rm XC}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}$ for small and large s. Therefore it neither behaves well under uniform scaling to high-density limit, nor describes the uniform electron gas as well as the LDA. The PBE [22] is a simpler functional and a smoother exchange-correlation potential, and so has better agreement with the LDA for small s limit. Generally GGAs produce poor results when the electron density has significant curvatures, which needs including the correction of the Laplacian of the density.

2.4 Pseudopotentials

In solid state systems, one assumes that only the valence electrons are important in determining the electronic structure of the system, while core electrons are taken to be localized at the atomic sites. As the wavefunctions of valence electrons are highly oscillatory within core region, they need to be expanded into a large number of plane waves. The aim of generating pseudopotentials is to produce smoother pseudo wavefunctions, which make computation efficient, and reproduce the all-electron behavior outside the core region.

Different schemes have been proposed to generate pseudopotentials, such as, norm-conserving pseudopotentials (NCPPs) [23, 30], ultra-soft pseudopotentials (USPPs) [24, 31] and projector augmented wave (PAW) potentials [25, 32]. All these three methods are in

common that they can reproduce the valence electronic properties of an all-electron calculation, and they are presently frozen core methods, i.e. the core electrons are pre-calculated in an atomic environment and kept frozen in the course of the remaining calculations.

To generate a pseudopotential, the following conditions must be fulfilled [30]. First, the pseudo wavefunction contains no nodes in the core region. Second, the pseudo and the all-electron wavefunction have the same eigenvalue, and agree beyond the core cutoff radius. The procedure of generating a pseudopotential is described as follows. One first calculates exact all-electron wavefunctions, $|\psi_{\ell m}\rangle$ from the Schrödinger equation,

$$[\hat{T} + V^{\mathrm{AE}}]|\psi_{\ell m}\rangle = \epsilon_{\ell m}|\psi_{\ell m}\rangle \,. \tag{2.8}$$

Then, choose a cutoff radius r_c and replace the exact wavefunction by a "soft" pseudo wavefunction with enforcing some continuity conditions. Beyond the "core region" of radius r_c, the pseudo-potentials and pseudo wavefunctions match with all-electron potentials and all-electron wavefunctions with identical eigenvalues. The pseudo wavefunctions $|\tilde{\psi}_{\ell m}\rangle$ contain no radial nodes within r_c. With the above constraints, one constructs pseudo wavefunction $|\tilde{\psi}_{\ell m}\rangle$, and pseudo-potential $\tilde{V}$ obeying the Kohn-Sham equation:

$$[\hat{T} + \tilde{V}]|\tilde{\psi}_{\ell m}\rangle = \epsilon_{\ell m}|\tilde{\psi}_{\ell m}\rangle \,. \tag{2.9}$$

2.4.1 Norm-conserving pseudopotentials (NCPPs)

The integrated charge inside radius r_c agrees for normalized pseudopotential and for all-electron wavefunction for any valence state,

$$\langle\tilde{\psi}_{\ell m}|\tilde{\psi}_{\ell m}\rangle_{r<r_c} = \langle\psi_{\ell m}|\psi_{\ell m}\rangle_{r<r_c} \,. \tag{2.10}$$

Many different schemes can be used to construct norm-conserving pseudopotentials, such as Hamann-Schlüter-Chiang [23], Troullier-Martins [33], etc. In our work, we use Troullier-Martins type pseudopotentials.

The pseudized valence state can be written as

$$\tilde{\psi}_{\ell m}(\mathbf{r}) = [\tilde{u}_\ell(r)/r]\, Y_{\ell m}(\theta, \phi)\,, \tag{2.11}$$

with $Y_{\ell m}(\theta, \phi) = P_\ell(cos(\theta))e^{im\phi}$. The radial part $\tilde{u}_\ell(r)$ for angular momentum ℓ is solved from the nonrelativistic Schrödinger equation:

$$\left[-\frac{1}{2}\frac{d^2}{dr^2} + \frac{\ell(\ell+1)}{2r^2} + \tilde{V}_\ell^{\rm scr}(r) - \tilde{\epsilon}_\ell\right] \tilde{u}_\ell(\tilde{\epsilon}_\ell, r) = 0\,. \tag{2.12}$$

Once the pseudo wavefunction is obtained, the pseudopotential can be solved from the inversion of the radial Schrödinger Equation, $\tilde{V}_\ell^{\rm scr}(r) = \tilde{\epsilon}_\ell - \frac{\ell(\ell+1)}{2r^2} + \frac{1}{2\tilde{u}_\ell(r)}\frac{d^2}{dr^2}\tilde{u}_\ell(r)$. To obtain a continuous pseudopotential which is regular at the origin, one additional constraint is introduced as, $\tilde{u}_\ell(r)$ is second differentiable, and behaves as $r^{\ell+1}$ near the origin. Furthermore, the transferability should be checked before extensive calculations. The norm-conserving constraints ensure the logarithmic derivatives of pseudo and all-electron agree around each reference energy ϵ_ℓ, and to first order in the energy, $\frac{d}{d\epsilon}\frac{d}{dr}\ln u_\ell(\epsilon, r)$. Also, pseudopotentials should reproduce the excitation energies for atoms.

In constructing the pseudopotentials, the exchange-correlation scheme should be determined as the one will be used for later poly-atomic system. The pseudopotentials are dependent on the angular momentum quantum number ℓ. One decomposes the semilocal potential operator into local part and a short-ranged nonlocal part (only for $\ell < \ell_{\rm max}$), $\hat{V}_{\rm SL} = V^{\rm loc}(\mathbf{r}) + \sum_{\ell=0}^{\ell_{\rm max}} \sum_m \delta V_\ell(r)|Y_{\ell m}\rangle\langle Y_{\ell m}|$.

Generally, the nonlocality of pseudopotential is restricted to the angular parts, so called semilocal. Kleinman and Bylander (KB) [34] make it possible to construct a fully non-

local separable form of pseudopotential for more efficient calculations. It can reduce the calculation and storage of matrix element $\langle\psi_i|\delta V_\ell^{\rm nl}|\psi_j\rangle$ from $mn(n+1)/2$ to mn, where m is the number of included k points, n is the number of plane wave expansion. The KB pseudopotential is given by

$$\hat{V}_{\rm SL}(\mathbf{r},\mathbf{r}') = V^{\rm loc}(\mathbf{r}) + \sum_{\ell m} \frac{|\delta V_\ell(r)\tilde{\psi}_{\ell m}(\mathbf{r})\rangle\langle\tilde{\psi}_{\ell m}(\mathbf{r}')\delta V_\ell(r')|}{\langle\tilde{\psi}_{\ell m}(\mathbf{r})|\delta V_\ell(\mathbf{r})|\tilde{\psi}_{\ell m}(\mathbf{r})\rangle}\,. \tag{2.13}$$

$\tilde{\psi}_{\ell m}(\mathbf{r})$ are pseudo wavefunctions defined in Eqn. (2.11). As the angular behavior is factored out, we look at the radial equation. The KB operator is defined as

$$\delta\hat{V}_\ell^{\rm KB} = \frac{|\delta V_\ell(r)\tilde{u}_\ell(r)\rangle\langle\tilde{u}_\ell(r')\delta V_\ell(r')|}{\langle\tilde{u}_\ell(r)|\delta V_\ell(r)|\tilde{u}_\ell(r)\rangle}\,, \tag{2.14}$$

with the corresponding KB eigenvector as $|\delta V_\ell(r)\tilde{u}_\ell(r)\rangle$, and the KB eigenenergy is $\epsilon_\ell^{\rm KB} = \frac{\int_0^{r^{\rm max}}[\delta V_\ell(r)]^2[\tilde{u}_\ell(r)]^2 dr}{\int_0^{r^{\rm max}}\delta V_\ell(r)[\tilde{u}_\ell(r)]^2 dr}$, which determines the strength of the nonlocality.

However KB pseudopotential may give rise to unphysical "ghost" state [35], which has lower energy than zero-node reference state. For example, a ghost state exists in p orbital for Ga pseudopotential. To avoid the ghost state, we can change the choice of local orbital or adjust the cutoff radii of the offending or local component in semilocal pseudopotential. In this case of Ga, we can change the local orbital from d orbital to s orbital, or properly choose cutoff radii.

Total energy

In the norm-conserving pseudopotential approximation, the standard form of a many-body system total energy, Eqn. (2.2), is rewritten as

$$\begin{aligned} E_{\rm tot} = &-\frac{1}{2}\sum_{n\mathbf{k}} f_{n\mathbf{k}} \int d\mathbf{r}\,\tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r})\nabla^2\tilde{\psi}_{n\mathbf{k}}(\mathbf{r}) + E_{\rm XC}[\rho^e(\mathbf{r})] + \sum_\mu E_\mu^{\rm nl} \\ &+ \int d\mathbf{r}\,\rho^e(\mathbf{r})\sum_\mu V_\mu^{\rm loc}(\mathbf{r}-\mathbf{R}_\mu) + \frac{1}{2}\int d\mathbf{r}\,\rho^e(\mathbf{r})V_{\rm H}(\mathbf{r}) + \sum_{\mu<\nu}\frac{Z_\mu Z_\nu}{R_{\mu\nu}}\,, \end{aligned} \tag{2.15}$$

where $\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})$ is the valence pseudo wavefunction for the n^{th} band, for wavevectors $\mathbf{k}$ within the first Brillouin zone. $\rho^e(\mathbf{r})$ is the pseudo valence electron density in NCPPs. The first term is the pseudo electron kinetic energy. The second term is the exchange-correlation energy $E_{\rm XC}[\rho^e(\mathbf{r})] = \int d\mathbf{r}\, \rho^e(\mathbf{r})\varepsilon_{\rm XC}(\rho^e(\mathbf{r}), |\nabla(\rho^e(\mathbf{r}))|)$, where $\varepsilon_{\rm XC}$ is the exchange-correlation energy per electron; in the local density (LDA) or a generalized gradient (GGA) approximation it is a function of the density or the density and its gradient. The fermion nature of many-body interacting electrons is approximated by this exchange-correlation potential. The third term is the energy due to the non-local part of the pseudopotential

$$E_\mu^{\rm nl} = \sum_{n\mathbf{k}} \sum_\ell \int d\mathbf{r}\, \tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r}) V^{\rm nl}_{\mu\ell}(|\mathbf{r} - \mathbf{R}_\mu|) \wp_\ell \tilde{\psi}_{n\mathbf{k}}(\mathbf{r})\,, \tag{2.16}$$

where $V^{\rm nl}_{\mu\ell}$ is the ℓth component of the non-local pseudopotential, with $\wp_\ell$ the projection operator on angular momentum ℓ. This term is site-localized (non-zero only within the core radius around a site) so that the total energy involves a sum over the sites μ at position $\mathbf{R}_\mu$. The last three terms of Eqn. (2.15) are the long-ranged Coulomb interactions. The fourth and fifth terms are the interaction of the electrons with the local ionic pseudopotential $V^{\rm loc}_\mu(\mathbf{r} - \mathbf{R}_\mu)$ and with themselves that can be written as one-half the interaction with the Hartree potential $V_{\rm H}(\mathbf{r}) = \int d\mathbf{r}' \frac{\rho^e(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$. The sixth term is the valence charge Z_μ of ion μ at position $\mathbf{R}_\mu$—the ion-ion Coulomb interaction energy that is the same as for point charges since the cores are assumed not to overlap. Finally, there can be non-linear core corrections not shown here, but which can be expressed in terms of $E_{\rm XC}$ involving a core density.

2.4.2 Ultrasoft pseudopotentials (USPPs)

Another accurate but smoother pseudopotential is called "ultrasoft" pseudopotentials (USPPs), developed by Vanderbilt in early 1990s [24], which does not force the "norm-conserving" condition, i.e.

$$\langle\tilde{\psi}_i|\tilde{\psi}_i\rangle_{r<r_c} \neq \langle\psi_i|\psi_i\rangle_{r<r_c}\,, \tag{2.17}$$

here composite index i denotes $\{\epsilon_i \ell m\}$, with the difference $q_{ij} = \langle\psi_i|\psi_i\rangle_{r<r_c} - \langle\tilde{\psi}_i|\tilde{\psi}_i\rangle_{r<r_c}$. For first-row elements and elements containing core d electrons, ultrasoft pseudopotentials are smoother than norm-conserving pseudopotentials, that is r_c's are chosen to be larger than for norm-conserving pseudopotentials. Thus smaller basis set will be needed in the calculation.

Since $V_{\rm AE} = V^{\rm loc}$ and $\tilde{\psi}_i = \psi_i$ outside the cutoff radius r_c, one can define a new function

$$|\tilde{\chi}_i\rangle = \left[\epsilon_i - \left(-\frac{1}{2}\nabla^2 + V^{\rm loc}\right)\right]|\tilde{\psi}_i\rangle\,, \tag{2.18}$$

which is non-zero only inside r_c. The semilocal potential operator can be written as

$$\hat{V}_{\rm SL}(\mathbf{r},\mathbf{r}') = V^{\rm loc} + V^{\rm nl} = V^{\rm loc} + \sum_i \frac{|\tilde{\chi}_i\rangle\langle\tilde{\chi}_i|}{\langle\tilde{\chi}_i|\tilde{\psi}_i\rangle}\,. \tag{2.19}$$

Define the matrix $B_{ij} = \langle\tilde{\psi}_i|\tilde{\chi}_j\rangle$ and the projector functions $|\beta_i\rangle = \sum_j (B^{-1})_{ji}|\tilde{\chi}_j\rangle$ dual to the $|\tilde{\psi}_i\rangle$, one can reach a generalized nonlocal pseudopotential operator

$$V^{\rm nl} = \sum_{ij} B_{ij}|\beta_i\rangle\langle\beta_j|\,. \tag{2.20}$$

The "norm-conserving" constraint can be removed as one introduces a new nonlocal pseudopotential operator [24]:

$$V^{\rm nl} = \sum_{ij} D_{ij}|\beta_i\rangle\langle\beta_j|\,, \tag{2.21}$$

where $D_{ij} = B_{ij} + \epsilon_j q_{ij}$; and an overlap operator

$$\hat{S} = I + \sum_{ij} q_{ij}|\beta_i\rangle\langle\beta_j|\,. \tag{2.22}$$

The orthonormality condition takes on a generalized form, $\langle\tilde{\psi}_i|\hat{S}|\tilde{\psi}_j\rangle = \delta_{ij}$.

Hamiltonian

The generalized Kohn-Sham equation can be written as

$$\left[-\frac{1}{2}\nabla^2 + \tilde{V}_{\text{eff}} + V^{\text{nl}}\right]\tilde{\psi}_{n\mathbf{k}} = \epsilon_{n\mathbf{k}}\hat{S}\tilde{\psi}_{n\mathbf{k}}\,, \tag{2.23}$$

with the effective potential

$$\tilde{V}_{\text{eff}} = V_{\text{H}}[\tilde{\rho} + \hat{\rho} + \tilde{\rho}_{Zc}] + V_{\text{XC}}[\tilde{\rho} + \hat{\rho} + \tilde{\rho}_c]\,. \tag{2.24}$$

$V_{\text{H}}[\tilde{\rho}_{Zc}]$ represents the local pseudopotential due to the pseudized core charge $\tilde{\rho}_{Zc}$. A pseudo electronic core density $\tilde{\rho}_c$ is included in exchange-correlation potential to improve the transferability of pseudopotentials [36].

Charge density

The deficit of valence electron density in the core region is recovered by a function

$$Q_{ij}(\mathbf{r}) = \psi_i^*(\mathbf{r})\psi_j(\mathbf{r}) - \tilde{\psi}_i^*(\mathbf{r})\tilde{\psi}_j(\mathbf{r})\,. \tag{2.25}$$

The total valence electron density $\rho^e(\mathbf{r})$ includes soft and hard parts. Soft part $\tilde{\rho}$ is the pseudo valence electron density. Hard part $\hat{\rho}$ is the compensation electron density.

$$\begin{aligned}\rho^e(\mathbf{r}) &= \tilde{\rho}(\mathbf{r}) + \hat{\rho}(\mathbf{r}) \\ &= \sum_{n\mathbf{k}} f_{n\mathbf{k}}\tilde{\psi}_{n\mathbf{k}}^*(\mathbf{r})\tilde{\psi}_{n\mathbf{k}}(\mathbf{r}) + \sum_{ij}\rho_{ij}Q_{ij}(\mathbf{r})\,,\end{aligned} \tag{2.26}$$

where matrix $\rho_{ij} = \sum_{n\mathbf{k}} f_{n\mathbf{k}}\langle\tilde{\psi}_{n\mathbf{k}}|\beta_i\rangle\langle\beta_j|\tilde{\psi}_{n\mathbf{k}}\rangle$.

Total energy

The total energy is given by

$$E_{\rm tot} = -\frac{1}{2}\sum_{n\mathbf{k}} f_{n\mathbf{k}} \int d\mathbf{r}\, \tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r})\nabla^2\tilde{\psi}_{n\mathbf{k}}(\mathbf{r}) + E_{\rm XC}[\rho^e(\mathbf{r})] + \sum_\mu E^{\rm nl}_\mu + \int d\mathbf{r}\, \rho^e(\mathbf{r}) \sum_\mu V^{\rm loc}_\mu(\mathbf{r}-\mathbf{R}_\mu) + \frac{1}{2}\int d\mathbf{r}\, \rho^e(\mathbf{r}) V_{\rm H}(\mathbf{r}) + \sum_{\mu<\nu} \frac{Z_\mu Z_\nu}{R_{\mu\nu}}\,. \tag{2.27}$$

The non-local part of pseudopotential energy is

$$E^{\rm nl}_\mu \;=\; \sum_{n\mathbf{k}} \int\, d\mathbf{r}\, \tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r}) \left(\sum_{ij} D^{\rm ion}_{ij} |\beta_i\rangle\langle\beta_j| \right) \tilde{\psi}_{n\mathbf{k}}(\mathbf{r})\,, \tag{2.28}$$

with $D^{\rm ion}_{ij} = D_{ij} - \int \tilde{V}_{\rm eff}(\mathbf{r}) Q_{ij}(\mathbf{r}) d\mathbf{r}$. The sum of eigenvalues is

$$\sum_{n\mathbf{k}} f_{n\mathbf{k}} \langle\tilde{\psi}_{n\mathbf{k}}|\hat{H}|\tilde{\psi}_{n\mathbf{k}}\rangle = \sum_{n\mathbf{k}} f_{n\mathbf{k}} \langle\tilde{\psi}_{n\mathbf{k}}| -\frac{1}{2}\nabla^2|\tilde{\psi}_{n\mathbf{k}}\rangle + \int \tilde{V}_{\rm eff}(\mathbf{r})(\tilde{\rho}(\mathbf{r}) + \hat{\rho}(\mathbf{r}))d\mathbf{r} + \sum_{ij} \rho_{ij} D^{\rm ion}_{ij}\,. \tag{2.29}$$

2.4.3 All-electron projector augmented wave method (PAWs)

The projector augmented wave method (PAW) is a combination of augmented plane wave method (APW) and pseudopotential approach. The ionic interaction with electrons are treated with a pseudopotential. The plane wave basis is a small set in Kohn-Sham equation.

Transformation theory and all-electron wavefunctions

A linear transformation $T = 1 + \sum_\mu \tilde{T}_\mu$ relates all-electron wavefunctions $|\psi_n\rangle$, strongly oscillating near the nucleus to soft pseudo wavefunctions $|\tilde{\psi}_n\rangle$. It has local contribution $\tilde{T}_\mu$ from non-overlapping spheres. In each sphere, an all-electron (AE) partial wave basis $|\phi_i\rangle$ is solved from the Schrödinger equation for isolated atom μ. Here, the index i refers to an ionic site index μ, angular momentum indices (ℓm) and an additional index ϵ that differentiates

partial waves with same angular momentum quantum numbers on the same site. The dual "soft" pseudo-partial wave $|\tilde{\phi}_i\rangle$ is built that matches to $|\phi_i\rangle$ at the sphere boundary, also outside the augmentation region. Within each augmentation region Ω_μ,

$$|\phi_i\rangle = (1+\tilde{T}_\mu)|\tilde{\phi}_i\rangle = |\tilde{\phi}_i\rangle + |\phi_i\rangle - |\tilde{\phi}_i\rangle,\ i \in R_\mu\,. \tag{2.30}$$

Therefore, $\tilde{T}_\mu|\tilde{\phi}_i\rangle = |\phi_i\rangle - |\tilde{\phi}_i\rangle$. Introduce smooth projector functions $|\tilde{p}_i\rangle$ dual to pseudo-partial waves, and fulfilling the orthogonality and completeness relations inside the augmentation spheres: $\langle\tilde{p}_i|\tilde{\phi}_i\rangle = \delta_{ij}$ for $i,j \in R_\mu$, $\sum_{i\in R_\mu}|\tilde{\phi}_i\rangle\langle\tilde{p}_i| = 1$. Therefore, within each Ω_μ,

$$\tilde{T}_\mu|\tilde{\psi}\rangle = \sum_{i\in R_\mu}\tilde{T}_\mu|\tilde{\phi}_i\rangle\langle\tilde{p}_i|\tilde{\psi}\rangle = \sum_{i\in R_\mu}(|\phi_i\rangle - |\tilde{\phi}_i\rangle)\langle\tilde{p}_i|\tilde{\psi}\rangle; \tag{2.31}$$

Outside all augmentation regions Ω_μ, $|\phi_i\rangle = |\tilde{\phi}_i\rangle$. At the end, we reach the linear transformation sum over all partial waves of all atoms:

$$T = 1 + \sum_i(|\phi_i\rangle - |\tilde{\phi}_i\rangle)\langle\tilde{p}_i|. \tag{2.32}$$

If the partial wave basis is complete, within each augmentation region Ω_μ, one would have $|\tilde{\psi}\rangle = \sum_{i\in R_\mu}|\tilde{\phi}_i\rangle C_i^\mu$, where $|\tilde{\phi}_i\rangle$ are non-orthogonal basis, and coefficients C_i^μ are scalar products $\langle\tilde{p}_i|\tilde{\psi}\rangle$ of projectors with pseudo wavefunctions. All-electron wavefunctions can then be written as:

$$\begin{aligned} |\psi_n\rangle = T|\tilde{\psi}_n\rangle =& |\tilde{\psi}_n\rangle + \sum_i(|\phi_i\rangle - |\tilde{\phi}_i\rangle)\langle\tilde{p}_i|\tilde{\psi}_n\rangle \\ =& |\tilde{\psi}_n\rangle + \sum_\mu(|\psi_\mu^1\rangle - |\tilde{\psi}_\mu^1\rangle)\,, \end{aligned} \tag{2.33}$$

with

$$
\begin{aligned}
|\psi_\mu^1\rangle &= \sum_{\epsilon\ell m} |\phi_{\epsilon\ell m}\rangle\langle\tilde{p}_{\epsilon\ell m}|\tilde{\psi}_n\rangle \,, \\
|\tilde{\psi}_\mu^1\rangle &= \sum_{\epsilon\ell m} |\tilde{\phi}_{\epsilon\ell m}\rangle\langle\tilde{p}_{\epsilon\ell m}|\tilde{\psi}_n\rangle \,.
\end{aligned}
\tag{2.34}
$$

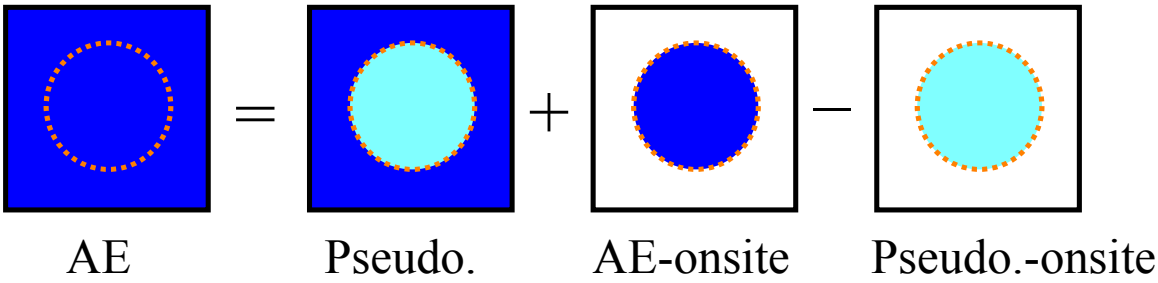

Figure 2.1: Schematic illustration of the PAW method. All physical quantities, such as all-electron wavefunctions, charge density, total energy, energy denties, etc., are written in three parts. All-electron and pseudo quantities match at and outside the augmentation spheres. Their difference within the sphere is compensated by the difference of all-electron-onsite and pseudo-onsite terms.

The "character" of an arbitrary pseudo wavefunction $|\tilde{\psi}_n\rangle$ at one site is given by $C_{\epsilon\ell m} = \langle\tilde{p}_{\epsilon\ell m}|\tilde{\psi}_n\rangle$. All partial waves and projector functions are determined before doing a calculation, and will not change during calculation. Pseudo wavefunction $\tilde{\psi}_n(\mathbf{r})$ is nodeless, and expanded in plane wave basis. $\psi_\mu^1(\mathbf{r})$ and $\tilde{\psi}_\mu^1(\mathbf{r})$ are all-electronic onsite and pseudo onsite wavefunction respectively, they both are written on a radial grid. Quantities with a tilde are obtained by pseudization, and a superscript 1 for quantities evaluated inside atom-centered spheres on a radial grid.

Operator expectation value

For an arbitrary operator A, the expectation value is

$$
\langle A\rangle = \sum_{n\mathbf{k}} f_{n\mathbf{k}}\langle\psi_{n\mathbf{k}}|A|\psi_{n\mathbf{k}}\rangle = \sum_{n\mathbf{k}} f_{n\mathbf{k}}\langle\tilde{\psi}_{n\mathbf{k}}|T^+AT|\tilde{\psi}_{n\mathbf{k}}\rangle \,. \tag{2.35}
$$

From the formalism of linear transformation T,

$$\begin{aligned}\tilde{A} &= T^{+}AT \\ &= \{1+\sum_i |\tilde{p}_i\rangle(\langle\phi_i| - \langle\tilde{\phi}_i|)\}A\{1+\sum_j(|\phi_j\rangle - |\tilde{\phi}_j\rangle)\langle\tilde{p}_j|\} \\ &= A+\sum_{i,j}|\tilde{p}_i\rangle(\langle\phi_i|A|\phi_j\rangle - \langle\tilde{\phi}_i|A|\tilde{\phi}_j\rangle)\langle\tilde{p}_j|\,. \end{aligned} \tag{2.36}$$

The above derivation uses two facts: 1. Within augmentation regions Ω_μ, $\sum_i |\tilde{\phi}_i\rangle\langle\tilde{p}_i| = 1$,which leads to $T^{+}AT = \{\sum_i |\tilde{p}_i\rangle\langle\phi_i|\}A\{\sum_j |\phi_j\rangle\langle\tilde{p}_j|\}$. 2. Outside all Ω_μ, $|\phi_i\rangle = |\tilde{\phi}_i\rangle$. Then, $T^{+}AT = A$. Therefore, the expectation value of any operator contains three parts in the PAW method, illustrated in Figure 2.1.

Partial waves and projectors

The PAW method depends upon the construction of partial waves and projector functions. The partial wave is the product of the radial function and spherical harmonics:

$$\phi_i(\mathbf{r}) = \frac{w_{\epsilon_i\ell_i}(r)}{r}Y_{\ell_i m_i}(\hat{\mathbf{r}})\tilde{\phi}_i(\mathbf{r}) = \frac{\tilde{w}_{\epsilon_i\ell_i}(r)}{r}Y_{\ell_i m_i}(\hat{\mathbf{r}})\,.$$

The standard scheme to construct pseudo partial wave follows RRKJ (spherical Bessel functions):

$$\tilde{\phi}_{\epsilon\ell}(r) = \begin{cases} \sum_{j=1}^{2}\alpha_j j_l(q_j r) & \text{if } r < r_c^\ell \\ \phi_{\epsilon\ell}(r) & \text{if } r > r_c^\ell \end{cases}$$

The projector function is written as

$$\tilde{p}_i(\mathbf{r}) = \frac{\tilde{p}_{\epsilon_i\ell_i}(r)}{r}Y_{\ell_i m_i}(\hat{\mathbf{r}})\,. \tag{2.37}$$

All-electron charge density

The on-site density matrix is

$$\rho_{ij} = \sum_{n\mathbf{k}} f_{n\mathbf{k}} \langle \tilde{\psi}_{n\mathbf{k}} | \tilde{p}_i \rangle \langle \tilde{p}_j | \tilde{\psi}_{n\mathbf{k}} \rangle \,, \tag{2.38}$$

which are the occupancies of each augmentation channel (i, j). The expectation value of charge density operator $|\mathbf{r}\rangle\langle\mathbf{r}|$ containing three terms

$$\begin{aligned} \rho(\mathbf{r}) &= \sum_{n\mathbf{k}} f_{n\mathbf{k}} \langle \psi_{n\mathbf{k}} | \mathbf{r} \rangle \langle \mathbf{r} | \psi_{n\mathbf{k}} \rangle \\ &= \sum_{n\mathbf{k}} f_{n\mathbf{k}} \langle \tilde{\psi}_{n\mathbf{k}} | T^{+} | \mathbf{r} \rangle \langle \mathbf{r} | T | \tilde{\psi}_{n\mathbf{k}} \rangle \\ &= \tilde{\rho}(\mathbf{r}) + \sum_{\mu} (\rho^1_{\mu}(\mathbf{r} - \mathbf{R}_{\mu}) - \tilde{\rho}^1_{\mu}(\mathbf{r} - \mathbf{R}_{\mu})) \,. \end{aligned} \tag{2.39}$$

The first term is the pseudo charge density evaluated from the pseudo wave functions on a plane wave grid, while second and third terms are evaluated on a spherical grid for each sphere contribution. Within a complete basis set, $\tilde{\rho}(\mathbf{r}) = \sum_{\mu} \tilde{\rho}^1_{\mu}(\mathbf{r} - \mathbf{R}_{\mu})$. In the following, they are discussed separately.

Pseudo soft charge density

The Bloch theorem states that in periodic crystal, the Kohn-Sham pseudo wavefunction is defined as

$$\tilde{\psi}_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r}) e^{i2\pi\mathbf{k}\cdot\mathbf{r}} = \sum_{\mathbf{G}} C_{n\,\mathbf{k}+\mathbf{G}} e^{i2\pi\,(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \,, \tag{2.40}$$

where $u_{n\mathbf{k}}(\mathbf{r})$ is a periodic function. $C_{n\,\mathbf{k}+\mathbf{G}}$ represent the coefficients in plane wave expansion. The corresponding pseudo soft charge density is given by

$$\tilde{\rho}(\mathbf{r}) = \sum_{n\,\mathbf{k}} f_{n\mathbf{k}} \tilde{\psi}^{*}_{n\mathbf{k}}(\mathbf{r}) \tilde{\psi}_{n\mathbf{k}}(\mathbf{r}) = \sum_{n\mathbf{k}} u^{*}_{n\mathbf{k}}(\mathbf{r}) u_{n\mathbf{k}}(\mathbf{r}) \,. \tag{2.41}$$

The scheme to construct the pseudo soft charge density according to given pseudo wavefunctions $\tilde{\psi}_{n\mathbf{k}}$ and electron occupation number $f_{n\mathbf{k}}$ is

$$\tilde{\psi}_{n\mathbf{k}}(\mathbf{G})|_{G_{\text{cut}}} \xrightarrow{\text{IFFT}} \tilde{\psi}_{n\mathbf{k}}(\mathbf{r}) \xrightarrow{\sum_{n\mathbf{k}} f_{n\mathbf{k}} \tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r}) \cdot \tilde{\psi}_{n\mathbf{k}}(\mathbf{r})} \tilde{\rho}(\mathbf{r}) \xrightarrow{\text{FFT}} \tilde{\rho}(\mathbf{G})|_{2G_{\text{cut}}} .$$

On-site charge density

On-site charge densities are localized around each atom, and can be evaluated in spherical grid accurately.

$$\begin{aligned}
\rho^1_\mu(\mathbf{r}) &= \sum_{i,j} \rho_{i,j} \langle \phi_i | \mathbf{r} \rangle \langle \mathbf{r} | \phi_j \rangle = \sum_{n\mathbf{k}ij} f_{n\mathbf{k}} \langle \tilde{\psi}_{n\mathbf{k}} | \tilde{p}_i \rangle \langle \tilde{p}_j | \tilde{\psi}_{n\mathbf{k}} \rangle \phi_i(\mathbf{r})^* \phi_j(\mathbf{r}) , \\
\tilde{\rho}^1_\mu(\mathbf{r}) &= \sum_{i,j} \rho_{i,j} \langle \tilde{\phi}_i | \mathbf{r} \rangle \langle \mathbf{r} | \tilde{\phi}_j \rangle = \sum_{n\mathbf{k}ij} f_{n\mathbf{k}} \langle \tilde{\psi}_{n\mathbf{k}} | \tilde{p}_i \rangle \langle \tilde{p}_j | \tilde{\psi}_{n\mathbf{k}} \rangle \tilde{\phi}_i(\mathbf{r})^* \tilde{\phi}_j(\mathbf{r}) .
\end{aligned} \tag{2.42}$$

In all-electron calculation, the wavefunction $\phi_{\epsilon\ell m}(\mathbf{r}) = \frac{w_{\epsilon\ell}(r)}{r} Y_{\ell m}(\theta, \varphi)$ refers to reference state with energy $\epsilon_{\epsilon\ell}$. Real spherical harmonics is used. On-site charge density

$$\begin{aligned}
\rho^1_\mu(\mathbf{r}) &= \sum_{ij} \rho_{ij} \langle \phi_i | \mathbf{r} \rangle \langle \mathbf{r} | \phi_j \rangle \\
&= \sum_{\epsilon_i \epsilon_j} \sum_{\ell_i \ell_j} \sum_{m_i m_j} \rho_{i,j} \frac{w_{\epsilon_i \ell_i}(r)}{r} \frac{w_{\epsilon_j \ell_j}(r)}{r} Y_{\ell_i m_i} Y_{\ell_j m_j} \\
&= \sum_{\epsilon_i \epsilon_j} \sum_{\ell_i \ell_j} \sum_{m_i m_j} \rho_{ij} \frac{1}{r^2} w_{\epsilon_i \ell_i}(r)\, w_{\epsilon_j \ell_j}(r) \sum_{L=|\ell_i - \ell_j|}^{\ell_i + \ell_j} G^{L\, m_i+m_j}_{\ell_i m_i, \ell_j m_j} Y^{L\, m_i+m_j}_{\ell_i \ell_j} .
\end{aligned} \tag{2.43}$$

$G^{L\, m_i+m_j}_{\ell_i m_i, \ell_j m_j}$ are the Gaunt coefficients. The scheme to construct the on-site charge density is

$$\begin{aligned}
\frac{1}{r^2} w_{\epsilon_i \ell_i}(r) w_{\epsilon_j \ell_j}(r)|_{\text{radial grid}} &\xrightarrow[\text{onto FFT grid}]{\text{map radial grid}} \frac{1}{r^2} w_{\epsilon_i \ell_i}(r) w_{\epsilon_j \ell_j}(r)|_{\text{FFT grid}} \\
&\xrightarrow[\cdot \rho_{i,j}]{\cdot Y_{\ell_i m_i}(\hat{\mathbf{r}}) Y_{\ell_j m_j}(\hat{\mathbf{r}})|_{\text{FFT grid}}} \rho^1_\mu(\mathbf{r})|_{\text{FFT grid}} .
\end{aligned}$$

Compensation charge

The total charge density includes pseudo charge density on FFT grid and on-site charge densities on spherical grid. One extra charge density, compensation charge is introduced in, so that Coulomb interaction of electrons would not include cross term between FFT grid and spherical grid. Charge density difference between all-electron and pseudo partial wave for each channel (i, j) is

$$Q_{ij}(\mathbf{r}) = \phi_i^*(\mathbf{r})\phi_j(\mathbf{r}) - \tilde{\phi}_i^*(\mathbf{r})\tilde{\phi}_j(\mathbf{r})\,. \tag{2.44}$$

Their moments are given by

$$q_{ij}^{LM} = \int_{\Omega_\mu} Q_{ij}(\mathbf{r})|\mathbf{r} - \mathbf{R}_\mu|^L Y_{LM}^*(\widehat{\mathbf{r} - \mathbf{R}_\mu})d\mathbf{r}\,. \tag{2.45}$$

A compensation charge is defined as

$$\hat{\rho} = \sum_{ij}\sum_{LM}(\int Q_{ij}(\mathbf{r})r^L d\mathbf{r})\rho_{ij}\, G_{ij}^{LM}\, Y_{LM}\, g_L(r)\,. \tag{2.46}$$

The compensation function g_L has the moment L equal to 1. [32] The total charge density contains all the following pieces:

$$\begin{aligned}
\rho_{\text{tot}} &= \rho + \rho_{Zc} \\
&= \tilde{\rho} + \rho^1 - \tilde{\rho}^1 + \rho_{Zc} \\
&= (\tilde{\rho} + \hat{\rho} + \tilde{\rho}_{Zc}) + (\rho^1 + \rho_{Zc}) - (\tilde{\rho}^1 + \hat{\rho} + \tilde{\rho}_{Zc})\,,
\end{aligned} \tag{2.47}$$

where the three terms are total pseudo charge density, total all-electron charge density, and total on-site pseudo charge density, respectively. They are all charge neutral terms. The ρ_{Zc} is the point charge density of the nuclei ρ_Z plus the frozen core AE charge density ρ_c. The compensation charge $\hat{\rho}$ is chosen so that $\tilde{\rho}^1 + \hat{\rho}$ has the same moments as ρ^1 within each sphere.

Total energy

Individual energy term generally contains three parts in the PAW method,

$$\begin{aligned} E_{\rm tot} =& \tilde{E} + \sum_{\mu}(E^1_\mu - \tilde{E}^1_\mu) \\ =& -\frac{1}{2}\sum_{n\mathbf{k}} f_{n\mathbf{k}} \int d\mathbf{r}\, \tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r})\nabla^2\tilde{\psi}_{n\mathbf{k}}(\mathbf{r}) + E_{\rm XC}[\tilde{\rho}+\hat{\rho}+\tilde{\rho}_c] + \sum_{\mu}(E^1_\mu - \tilde{E}^1_\mu) \\ &+ E_{\rm H}[\tilde{\rho}+\hat{\rho}] + \int V_{\rm H}[\tilde{\rho}_{Zc}](\tilde{\rho}+\hat{\rho})d\mathbf{r} + \sum_{\mu<\nu}\frac{Z_\mu Z_\nu}{R_{\mu\nu}}\,. \end{aligned} \tag{2.48}$$

$$\begin{aligned} E^1_\mu =& \sum_{ij}\rho_{ij}\langle\phi_i| - \frac{1}{2}\nabla^2|\phi_j\rangle + \overline{E_{\rm XC}[\rho^1+\rho_c]} + \overline{E_{\rm H}[\rho^1]} + \int V_{\rm H}[\rho_{Zc}](\rho^1)d\mathbf{r}\,, \\ \tilde{E}^1_\mu =& \sum_{ij}\rho_{ij}\langle\tilde{\phi}_i| - \frac{1}{2}\nabla^2|\tilde{\phi}_j\rangle + \overline{E_{\rm XC}[\tilde{\rho}^1+\hat{\rho}+\tilde{\rho}_c]} + \overline{E_{\rm H}[\tilde{\rho}^1+\hat{\rho}]} + \int V_{\rm H}[\tilde{\rho}_{Zc}](\tilde{\rho}^1+\hat{\rho})d\mathbf{r}\,. \end{aligned} \tag{2.49}$$

The smoothly pseudized part $\tilde{E}$ is evaluated on plane waves grid. On-site energies E^1_μ and $\tilde{E}^1_\mu$ are evaluated on spherical grid from each atomic sphere contribution. There is no cross term between the planewaves and the radial grids. In band-structure calculation, the total energy is evaluated as:

$$E = \sum_{n\mathbf{k}} f_{n\mathbf{k}}\langle\tilde{\psi}_{n\mathbf{k}}|\hat{H}|\tilde{\psi}_{n\mathbf{k}}\rangle + \tilde{E}_{\rm dc} + E^1_{\rm dc} - \tilde{E}^1_{\rm dc} + \sum_{\mu<\nu}\frac{Z_\mu Z_\nu}{R_{\mu\nu}}\,, \tag{2.50}$$

with

$$\begin{aligned} \tilde{E}_{\rm dc} =& -E_{\rm H}[\tilde{\rho}+\hat{\rho}] + E_{\rm XC}[\tilde{\rho}+\hat{\rho}+\tilde{\rho}_c] - \int V_{\rm XC}[\tilde{\rho}+\hat{\rho}+\tilde{\rho}_c](\tilde{\rho}+\hat{\rho})\,d\mathbf{r}\,, \\ E^1_{\rm dc} =& -\overline{E_{\rm H}[\rho^1]} + \overline{E_{\rm XC}[\rho^1+\rho_c]} - \int V_{\rm XC}[\rho^1+\rho_c](\rho^1)\,d\mathbf{r}\,, \\ \tilde{E}^1_{\rm dc} =& -\overline{E_{\rm H}[\tilde{\rho}^1+\hat{\rho}]} + \overline{E_{\rm XC}[\tilde{\rho}^1+\hat{\rho}+\tilde{\rho}_c]} - \int V_{\rm XC}[\tilde{\rho}^1+\hat{\rho}+\tilde{\rho}_c](\tilde{\rho}^1+\hat{\rho})\,d\mathbf{r}\,. \end{aligned} \tag{2.51}$$

Hamiltonian

Introducing the overlap operator

$$\hat{S} = T^{+}T = I + \sum_{ij} |\tilde{p}_i\rangle(\langle\phi_i|\phi_j\rangle - \langle\tilde{\phi}_i|\tilde{\phi}_j\rangle)\langle\tilde{p}_j| \,, \tag{2.52}$$

which fulfills the orthonormality condition on a generalized form $\langle\tilde{\psi}_n|\hat{S}|\tilde{\psi}_m\rangle = \delta_{nm}$, the generalized Kohn-Sham equation can be written as

$$\hat{H}\tilde{\psi}_{n\mathbf{k}} = \epsilon_{n\mathbf{k}}\hat{S}\tilde{\psi}_{n\mathbf{k}} \,. \tag{2.53}$$

Hamiltonian operator for the PAW method is the functional derivative of total energy (Eqn. (2.48)) with respect to the pseudo electron density operator $\tilde{\rho} = \sum_{n\mathbf{k}} f_{n\mathbf{k}}|\tilde{\psi}_{n\mathbf{k}}\rangle\langle\tilde{\psi}_{n\mathbf{k}}|$. With defining $\tilde{V}_{\text{eff}} = V_{\text{H}}[\tilde{\rho} + \hat{\rho} + \tilde{\rho}_{Zc}] + V_{\text{XC}}[\tilde{\rho} + \hat{\rho} + \tilde{\rho}_c]$,

$$\hat{H} = \frac{dE}{d\tilde{\rho}} = -\frac{1}{2}\nabla^2 + \tilde{V}_{\text{eff}} + \sum_{ij} |\tilde{p}_i\rangle(\hat{D}_{ij} + D^1_{ij} - \tilde{D}^1_{ij})\langle\tilde{p}_j|. \tag{2.54}$$

Similarly define $V^1_{\text{eff}} = V_{\text{H}}[\rho^1 + \rho_{Zc}] + V_{\text{XC}}[\rho^1 + \rho_c]$ and $\tilde{V}^1_{\text{eff}} = V_{\text{H}}[\tilde{\rho}^1 + \hat{\rho} + \tilde{\rho}_{Zc}] + V_{\text{XC}}[\tilde{\rho}^1 + \hat{\rho} + \tilde{\rho}_c]$, three D_{ij} terms are expressed as below

$$\begin{aligned}
\hat{D}_{ij} =& \sum_{\ell m} \int \tilde{V}_{\text{eff}}(\mathbf{r})\hat{Q}^{\ell m}_{ij}(\mathbf{r})\, d\mathbf{r} \,, \\
D^1_{ij} =& \langle\phi_i| - \frac{1}{2}\nabla^2 + V^1_{\text{eff}}|\phi_j\rangle \,, \\
\tilde{D}^1_{ij} =& \langle\tilde{\phi}_i| - \frac{1}{2}\nabla^2 + \tilde{V}^1_{\text{eff}}|\tilde{\phi}_j\rangle + \sum_{\ell m} \int \tilde{V}^1_{\text{eff}}(\mathbf{r})\hat{Q}^{\ell m}_{ij}(\mathbf{r})\, d\mathbf{r} \,.
\end{aligned} \tag{2.55}$$

D^1_{ij} and $\tilde{D}^1_{ij}$ are evaluated on a radial grid within each augmentation sphere. They are given as two values for each (i, j) pair. The sum of eigenvalues can be written as

$$\sum_{n\mathbf{k}} f_{n\mathbf{k}} \langle \tilde{\psi}_{n\mathbf{k}} | \hat{H} | \tilde{\psi}_{n\mathbf{k}} \rangle = \sum_{n\mathbf{k}} f_{n\mathbf{k}} \langle \tilde{\psi}_{n\mathbf{k}} | -\frac{1}{2}\nabla^2 | \tilde{\psi}_{n\mathbf{k}} \rangle + \int \tilde{V}_{\text{eff}} \tilde{\rho}(\mathbf{r})\, d\mathbf{r} + \sum_{ij} \rho_{ij} D_{ij}\,, \tag{2.56}$$

with the last term calculated as $\int \tilde{V}_{\text{eff}}(\mathbf{r}) \hat{\rho}(\mathbf{r}) d\mathbf{r} + \sum_{ij} \rho_{ij} (D^1_{ij} - \tilde{D}^1_{ij})$.

Chapter 3

Energy density method

The total energy is one of the most important quantities in a solid state system, as it can be used to describe the stable structure and other equilibrium properties. However, sometime we are interested in determining energies from only part of the system, not only the total energy. For example, the surface energy of a given facet of a crystal is meaningful for predicting the equilibrium crystal shape and preferred crystal growth directions, and should depend only upon the properties of the surface.

In general, surface energies can be determined unambiguously using two conventional total energy calculations, one bulk calculation and one slab calculation when the two surfaces of the slab are crystallographically identical, e.g. the (110) surface of a zinc-blende semiconductor. For some surfaces, such as the (111) and $(\bar{1}\bar{1}\bar{1})$ surfaces of a zinc-blende semiconductor GaAs, it is impossible to calculate absolute surface energies only from total energy calculations for geometrical reason. To solve this problem, Chetty and Martin [15, 16] proposed an energy density method, in which they introduced the concepts of the energy density, and applied the Wigner-Seitz cells to integrate the local energy density over specific regions V for a high-symmetry crystal. $\int_V e(\mathbf{r}) \rightarrow E$ gives the energy associated with that region. Although the energy density function is not well-defined, they have shown that integrals over certain volumes (multiply of Wigner-Seitz cells) such as surface regions or interfaces, are well-defined and useful. Therefore, the surface energy of either slab surface can be integrated independently.

The energy density formula derived by Chetty and Martin, and later used by others is a reciprocal-space expression for norm-conserving pseudopotential (NCPPs) with local den-

sity approximation (LDA), where the ion-ion interaction is treated in a similar manner as the Ewald sum [37] (see Appendix (C)). In this chapter, I reformulate the expression of real-space energy density for the projector augmented wave (PAW) method and the pseudopotential methods in the same manner. After introducing a fictitious ionic charge density, all electrostatic interactions are grouped together in real space. I implement the energy density method in the Vienna ab initio simulation package (VASP) [38]. Although the real space energy density contains gauge dependent terms, as pointed out by Chetty and Martin one can employ a supercell composed of an integral number of Wigner-Seitz cells to resolve this problem. See also Rapcewicz *et al.* [39] for constructing a Voronoi polyhedron for each comprised atom. Specifically to this work, I apply the Bader volume [17] bounded by zero flux surface of the gradient of electron density to evaluate unique atomic kinetic energy, and construct a new charge neutral volume bounded by zero flux surface of total electrostatic field to evaluate unique classical Coulomb energy. Hence, the gauge dependent problem can be solved automatically.

The standard form of total energy in DFT is given in Eqn. (2.2). The energy density corresponding to each term in Eqn. (2.2) can be re-expressed as

$$e(\mathbf{r}) = t^{\mathrm{AE}}(\mathbf{r}) + e_{\mathrm{CC}}^{\mathrm{AE}}(\mathbf{r}) + e_{\mathrm{XC}}^{\mathrm{AE}}(\mathbf{r})\,. \tag{3.1}$$

Here it is crucial to group the three terms E_{H}, $\int d\mathbf{r}\rho^{e}(\mathbf{r})V_{\mathrm{ext}}(\mathbf{r})$, and E_{II} terms, as a total classical Coulomb energy $E_{\mathrm{CC}}^{\mathrm{AE}}$, because individual terms would diverge under the periodic boundary condition. Since the core states are tightly bonded and do not change significantly from one atom to another atom, it is advantageous to isolate such terms by introducing the on-site energy, and the total energy density can be written as

$$e(\mathbf{r}) = t(\mathbf{r}) + e_{\mathrm{CC}}(\mathbf{r}) + e_{\mathrm{XC}}(\mathbf{r}) + E_{\text{on-site}}\delta(\mathbf{r}-\mathbf{R}_{\mu})\,. \tag{3.2}$$

The kinetic energy density is known to be gauge-dependent, which we will discuss in Section 3.1. The classical Coulomb energy density contains all interactions among electrons and ions. In Chetty and Martin's work, a background charge composed of Gaussian distributions at each ionic site is included in a similar manner as in the Ewald Sum, which we will review in Section 3.2.

In Section 3.1, we first review the kinetic energy density with a special focus on the PAW method. Then we discuss the gauge-dependent long-range classical Coulomb energy density in Section 3.2, a well-defined exchange-correlation energy density in Section 3.3, and all the rest short-range terms grouped in an ion-based non-local energy in Section 3.4. Finally, we summarize our reformulated expression of energy density in Section 3.5.

3.1 Kinetic energy density

The kinetic energy density of independent fermions $t(\mathbf{r})$ is not unique, which has the following two different formulae

$$\begin{aligned} t^{(a)}(\mathbf{r}) &= -\frac{1}{2}\sum_{n\mathbf{k}} f_{n\mathbf{k}}\tilde{\psi}^{*}_{n\mathbf{k}}(\mathbf{r})\nabla^2\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})\,, \\ t^{(s)}(\mathbf{r}) &= \frac{1}{2}\sum_{n\mathbf{k}} f_{n\mathbf{k}}|\nabla\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})|^2\,. \end{aligned} \tag{3.3}$$

The former one is real-space asymmetric, while the later one is real-space symmetric. The integral of the two forms of kinetic energy density is equal either when the boundary terms cancel out in the periodic boundary condition or when the boundary terms vanish for the bounded states with wavefunctions $\tilde{\psi}_{n\mathbf{k}}$ vanishing at the boundary. The difference between the asymmetric and the symmetric kinetic energy density is gauge dependent term propor-

tional to the Laplacian of pseudo electron density,

$$\begin{aligned} t^{(a)}(\mathbf{r}) - t^{(s)}(\mathbf{r}) &= -\frac{1}{2}\sum_{n\mathbf{k}} f_{n\mathbf{k}} \left[\frac{1}{2}\tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r})\nabla^2\tilde{\psi}_{n\mathbf{k}}(\mathbf{r}) + \frac{1}{2}\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})\nabla^2\tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r}) + \nabla\tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r})\cdot\nabla\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})\right] \\ &= -\frac{1}{4}\sum_{n\mathbf{k}} f_{n\mathbf{k}} \left[\nabla\cdot(\tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r})\nabla\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})) + \nabla\cdot(\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})\nabla\tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r}))\right] \\ &= -\frac{1}{4}\nabla^2\tilde{\rho}(\mathbf{r})\,. \end{aligned} \tag{3.4}$$

In this work, we will integrate over a discrete set of atom-centered volumes where the gauge-dependent integrals also vanishes—hence, uniquely defined kinetic energies for atoms in a condensed system.

For a continuous wavefunction, its first-order derivative may be discontinuous thus the antisymmetric form of the kinetic energy density may have singularities. For this reason, Chetty and Martin chose the symmetric form for the kinetic energy density as it directly enters the variational principle for deriving the Schrödinger's equation [40] and hence is a more fundamental quantity. However, the kinetic energy density is unique except for terms proportional to the Laplacian of pseudo electron density; if we integrate over volumes where the gauge-dependent term of Eqn. (3.4), then either form of Eqn. (3.3) gives the same kinetic energy. For a planewave basis, the asymmetric kinetic energy density is well-defined everywhere—i.e., there are no cusps in the wavefunction gradient—and is computationally less demanding to calculate. In the PAW method, the total kinetic energy density contains three terms

$$t^{(a)}(\mathbf{r}) = \tilde{t}^{(a)}(\mathbf{r}) + t^{1(a)}(\mathbf{r}) - \tilde{t}^{1(a)}(\mathbf{r})\,. \tag{3.5}$$

The first term, a pseudo-kinetic energy density $\tilde{t}^{(a)}(\mathbf{r})$ is also the first term in Eqn. (2.48) which can be expressed as Eqn. (3.3) by using the pseudo wavefunction. The last two terms, on-site kinetic energies $t^{1(a)}(\mathbf{r})$ and $\tilde{t}^{1(a)}(\mathbf{r})$ are first terms in Eqn. (2.49), and are included in the short-ranged on-site energy $E_{\text{on-site}}\,\delta(\mathbf{r}-\mathbf{R}_\mu)$. Practically working on a planewave basis, one first calculates the Laplacian of the pseudo wavefunction in reciprocal space, then

inverse Fourier transform the result back into real space. After that, one multiplies it by the pseudo wavefunction and the occupation number to evaluate the asymmetric form of the soft kinetic energy density. Schematically, this procedure is

$$\tilde{\psi}_{n\mathbf{k}}(\mathbf{G}) \xrightarrow{\times|\mathbf{k}+\mathbf{G}|^2} |\mathbf{k}+\mathbf{G}|^2\tilde{\psi}_{n\mathbf{k}}(\mathbf{G}) \xrightarrow{\text{IFFT}} \nabla^2\tilde{\psi}_{n\mathbf{k}}(\mathbf{r}) \xrightarrow{\sum_{n\mathbf{k}} f_{n\mathbf{k}}\tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r})\nabla^2\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})} \tilde{t}^{(a)}(\mathbf{r})\,.$$

The gauge dependent term of the pseudo-kinetic energy density is proportional to the Laplacian of the pseudo charge density

$$\tilde{\rho}(\mathbf{G}) \xrightarrow{\times|\mathbf{k}+\mathbf{G}|^2} |\mathbf{k}+\mathbf{G}|^2\tilde{\rho}(\mathbf{G}) \xrightarrow{\text{FFT}} \nabla^2\tilde{\rho}(\mathbf{r}) \xrightarrow{\times\frac{1}{4}} \tilde{t}^{(a)}(\mathbf{r}) - \tilde{t}^{(s)}(\mathbf{r})\,.$$

3.2 Classical Coulomb energy density

The total classical Coulomb energy of a system with electrons and nuclei can be written as

$$E_{\text{CC}} = \frac{1}{2}\int d\mathbf{r}\int d\mathbf{r}'\frac{\rho^e(\mathbf{r})\rho^e(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \int d\mathbf{r}\rho^e(\mathbf{r})\sum_\mu V^{\text{loc}}_\mu(\mathbf{r}) + \sum_{\mu<\nu}\frac{Z_\mu Z_\nu}{R_{\mu\nu}}\,, \tag{3.6}$$

where μ and ν are representing different nuclei and $R_{\mu\nu}$ is the distance between two nuclei, and $\rho^e(\mathbf{r})$ represents the pseudoelectron density in NCPPs, and the sum of pseudoelectron density $\tilde{\rho}(\mathbf{r})$ and compensation charge $\hat{\rho}(\mathbf{r})$ for the USPPs and the PAW method. For a neutral system in a periodic boundary condition, individual terms of the above equation diverge. However, we can regroup those three terms in a different way such that the summed Coulomb energy is not divergent. There are various ways to calculate the electrostatic energy. [41] In the Ewald method (see Appendix (C)), one can first add and then subtract an ionic charge density with a Gaussian-like distribution localized around each atom

$$\rho^{\text{ion}}(\mathbf{r}) = \sum_\mu \rho_\mu(\mathbf{r}-\mathbf{r}_\mu) = \sum_\mu \frac{Z_\mu}{(R^{c,\mu})^3\,\pi^{3/2}}\exp\left[-\frac{|\mathbf{r}-\mathbf{r}_\mu|^2}{(R^{c,\mu})^2}\right]\,. \tag{3.7}$$

Properly choosing a Gaussian width $R^{c,\mu}$'s value, the effect of real space summation can be negligibly small. There are two terms left in the ion-ion interactions. One is the Hartree energy in terms of the Gaussian charge density in reciprocal space. The other one is the self energy term, which can be easily calculated as a function of the charge and the Gaussian width of each ion. To obtain the total Coulomb energy of both electrons and ions, one can group the electron-electron and ion-ion interactions by defining a total charge density $\rho^{\rm tot}$ as a sum of the valence electron charge density ρ^e and the fictitious ionic charge density $\rho^{\rm ion}$: $\rho^{\rm tot}(\mathbf{r}) = \rho^e(\mathbf{r}) + \rho^{\rm ion}(\mathbf{r})$. After some manipulation, the total classical Coulomb energy is re-expressed as

$$\begin{aligned} E_{\rm CC} &= \frac{1}{2}\int d\mathbf{r}\int d\mathbf{r}'\frac{\rho^{\rm tot}(\mathbf{r})\rho^{\rm tot}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \int d\mathbf{r}\,\rho^e\,(V^{\rm loc}-V^{\rm ion}) \\ &+ \frac{1}{2}\sum_{\mu\neq\nu}\frac{Z_\mu Z_\nu}{R_{\mu\nu}}\,{\rm erfc}\left[\frac{R_{\mu\nu}}{\sqrt{(R^{c,\mu})^2+(R^{c,\nu})^2}}\right] - \frac{1}{\sqrt{2\pi}}\sum_\mu\frac{Z_\mu^2}{R^{c,\mu}}\,. \end{aligned} \tag{3.8}$$

Here, the first term is the Hartree energy in term of the total neutral pseudo charge density. In the second term, the term, $V^{\rm ion} = \int d\mathbf{r}'\rho^{\rm ion}(\mathbf{r}')/|\mathbf{r}-\mathbf{r}'|$, represents the interaction between the electron and the Gaussian-broadened ions, and is subtracted from the local pseudopotential energy. The third term depends on overlaps of charge densities of ions located at different sites and should in principle be small. Due to the property of the erfc function, one can show that by choosing small Gaussian widths $R^{c,\mu}$ this overlap energy is negligibly small. The last term describes the self energy of pseudo-ions, which can be treated as a constant for each species of ions. In the classical electromagnetic theory, the Coulomb interaction can be expressed as the Maxwell energy in terms of the electric field or the electric potential, $E^{\rm Maxwell} = \int d\mathbf{r}\,|E(\mathbf{r})|^2/(8\pi) = \int d\mathbf{r}\,|\nabla V(\mathbf{r})|^2/(8\pi)$. While the Hartree energy can easily be transformed into the Maxwell energy form, other terms in Eqn. (3.8) require special cares. In the next section, we use methods that involve only smooth densities for each ion to construct expressions for the Coulomb energy that are expressed only in real space.

3.2.1 Smeared ions

In the present approach, we introduce a fictitious localized charge distribution $\rho_\mu^{\rm loc}$, which gives rise to a local pseudopotential $V_\mu^{\rm loc}$ (c.f. Section F.3 of [41]) for ion μ as

$$\rho_\mu^{\rm loc}(\mathbf{r}-\mathbf{R}_\mu) = -\frac{1}{4\pi}\nabla^2 V_\mu^{\rm loc}(\mathbf{r}-\mathbf{R}_\mu)\,. \tag{3.9}$$

The Coulomb interaction energy between two ions μ and ν is then given by

$$E_{\mu\nu}^{\rm loc}(|\mathbf{R}_{\mu\nu}|) = \frac{Z_\mu Z_\nu}{R_{\mu\nu}} = \int d\mathbf{r}\rho_\mu^{\rm loc}(\mathbf{r}-\mathbf{R}_\mu)V_\nu^{\rm loc}(\mathbf{r}-\mathbf{R}_\nu)\,, \tag{3.10}$$

and the self energy on each ion is defined as

$$E_\mu^{\rm self} = \frac{1}{2}\int d\mathbf{r}\rho_\mu^{\rm loc}(\mathbf{r})V_\mu^{\rm loc}(\mathbf{r})\,. \tag{3.11}$$

The total classical Coulomb energy of a system with electrons and nuclei can be simply written as

$$E_{\rm CC} = \int d\mathbf{r}\,\frac{1}{8\pi}|\nabla V^{\rm tot}(\mathbf{r})|^2 - \sum_\mu E_\mu^{\rm self} \tag{3.12}$$

with total classical Coulomb potential $V^{\rm tot}(\mathbf{r}) = V_{\rm H}(\mathbf{r}) + V^{\rm loc}(\mathbf{r})$. The Hartree, local, and ion-ion interaction terms (last three terms of Eqn. (2.48)) are combined into the classical Coulomb term.

Instead of using the Maxwell energy expression only for the Hartree energy in terms of the neutral pseudo charge, here we apply it onto the total electrostatic component of the energy density. As a fictitious charge density defined in Eqn. (3.9), the total neutral charge density $\rho^{\rm tot}(\mathbf{r}) = \rho^e(\mathbf{r}) + \rho^{\rm loc}(\mathbf{r})$. With this definition, the Maxwell form of the classical Coulomb energy density can be written as

$$e_{\rm CC}^{\rm Maxwell}(\mathbf{r}) = \frac{1}{8\pi}\,|\nabla V^{\rm tot}(\mathbf{r})|^2\,. \tag{3.13}$$

Similar to the kinetic energy density, the classical Coulomb energy density is not unique either because of gauge dependence. It can also be expressed as an asymmetric form,

$$e_{\mathrm{CC}}^{(a)}(\mathbf{r}) = -\frac{1}{8\pi} V^{\mathrm{tot}}(\mathbf{r}) \nabla^2 V^{\mathrm{tot}}(\mathbf{r}) = \frac{1}{2} V^{\mathrm{tot}}(\mathbf{r}) \rho^{\mathrm{tot}}(\mathbf{r}) \,, \tag{3.14}$$

and the gauge-dependent term is the difference of Eqn. (3.13) and Eqn. (3.14),

$$e_{\mathrm{CC}}^{(a)}(\mathbf{r}) - e_{\mathrm{CC}}^{\mathrm{Maxwell}}(\mathbf{r}) = -\frac{1}{8\pi} \nabla \cdot \left[V^{\mathrm{tot}}(\mathbf{r}) \nabla V^{\mathrm{tot}}(\mathbf{r}) \right] \,. \tag{3.15}$$

In a way analogous to the kinetic energy density, we can obtain a gauge-independent classical Coulomb energy as an integral over any volume bounded by a zero flux surface of the gradient of the total Coulomb potential.

Specifically for the PAW method and after introducing the soft compensation charge $\hat{n}$, the Hartree energy contains three terms:

$$\begin{aligned} E_{\mathrm{H}} =& \tilde{E}_{\mathrm{H}} + (E_{\mathrm{H}}^1 - \tilde{E}_{\mathrm{H}}^1) \\ =& E_{\mathrm{H}}[\tilde{\rho} + \hat{\rho}] + \sum_{\mu} \overline{E_{\mathrm{H}}}[\rho^1] - \sum_{\mu} \overline{E_{\mathrm{H}}[\tilde{\rho}^1 + \hat{\rho}]} \,. \end{aligned} \tag{3.16}$$

The first term is related to the soft valence charge density and the soft compensation charge density, and is included in the classical Coulomb energy density. The last two short-ranged terms are related to the short-ranged on-site energy $E_{\text{on-site}}\, \delta(\mathbf{r} - \mathbf{R}_\mu)$, as are the electron-ion interactions.

3.2.2 Model charge

In practice, the local charge density due to local pseudopotential can vary rapidly (c.f. Figure 3.1), which causes numerical errors in a real calculation. To improve our numerical accuracy, we then introduce a model charge density $\rho^{\mathrm{model}}(\mathbf{r})$ to solve this problem. The model charge density is chosen as obeying following constraints: a spherically symmetric

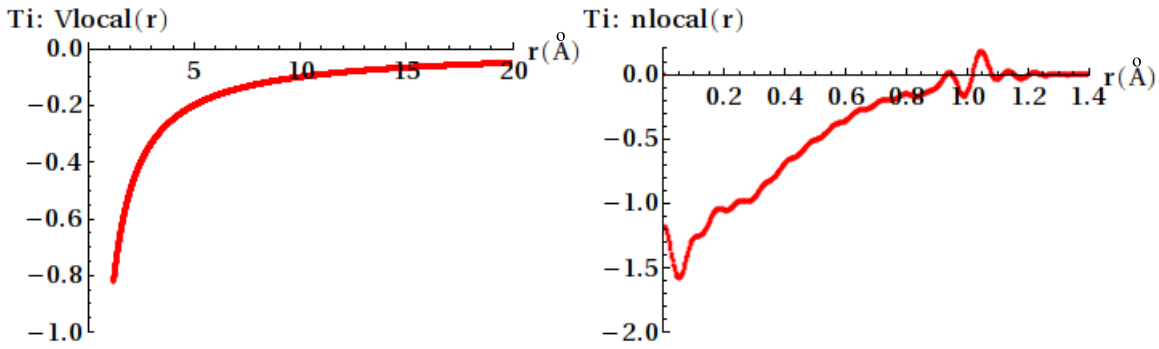

Figure 3.1: Local pseudopotential (the left panel) and change density (the right panel) for PAW Ti. The PAW-GGA potential for Ti has a cutoff radius 1.22 Å; the charge density—given by the Laplacian of the potential—can have short-wavelength oscillations that are well-represented on a radial grid, but poorly represented on a regular Cartesian grid.

functional, which is centered at each ion, zero beyond the cutoff radius of local pseudopotential, and normalized as negative to the local charge density within the cutoff radius. The total charge density can be rewritten as

$$\begin{aligned}\rho^{\mathrm{tot}}(\mathbf{r}) =& \rho^{\mathrm{loc}}(\mathbf{r}) + \rho^{\mathrm{model}}(\mathbf{r}) + \rho^{e}(\mathbf{r}) - \rho^{\mathrm{model}}(\mathbf{r}) \\ =& \sum_{\mu} \left[\rho_{\mu}^{\mathrm{loc}}(\mathbf{r}) + \rho_{\mu}^{\mathrm{model}}(\mathbf{r})\right] + \delta\rho(\mathbf{r})\,,\end{aligned} \tag{3.17}$$

where $\rho_{\mu}^{\mathrm{loc}}(\mathbf{r}) + \rho_{\mu}^{\mathrm{model}}(\mathbf{r})$ is a neutral and spherical charge density for each ion; $\delta\rho(\mathbf{r})$ is the difference between the valence electronic charge density and the model charge density. The asymmetric form of the classical Coulomb energy is then regrouped as

$$E_{\mathrm{CC}}[\rho^{\mathrm{tot}}] = E_{\mathrm{CC}}[\rho^{\mathrm{loc}} + \rho^{\mathrm{model}}] + \int d\mathbf{r}\,(V^{\mathrm{loc}} + V^{\mathrm{model}})\delta\rho + \frac{1}{2}\int d\mathbf{r}\,\delta V_{\mathrm{H}}[\delta\rho]\,\delta\rho - \sum_{\mu} E_{\mu}^{\mathrm{self}}\,. \tag{3.18}$$

The first term in Eqn. (3.18) is

$$E_{\mathrm{CC}}[\rho^{\mathrm{loc}} + \rho^{\mathrm{model}}] = \sum_{\mu<\nu} E_{\mu\nu}^{\mathrm{loc+model}}(|\mathbf{R}_{\mu\nu}|) + \sum_{\mu} E_{\mu}^{\mathrm{loc+model}}\,. \tag{3.19}$$

The electronic interaction between two neutral atoms is zero when there is no charge overlap, since all moments are zero for the spherical charge distributions. Therefore, the first term in Eqn. (3.19) is zero. Combining the second term with the self energy in Eqn. (3.11), we have

$$E_\mu^{\text{loc+model}} - E_\mu^{\text{self}} = \frac{1}{2}\int d\mathbf{r}\, V_\mu^{\text{model}}(\mathbf{r})\rho_\mu^{\text{model}}(\mathbf{r}) + \int d\mathbf{r}\, V_\mu^{\text{loc}}(\mathbf{r})\rho_\mu^{\text{model}}(\mathbf{r})\,, \tag{3.20}$$

which is a constant for each species of ions and can be canceled when studying the defect energies. Neglecting this constant term, the asymmetric form of the classical Coulomb energy density in Eqn. (3.18) is

$$e_{\text{CC}}(\mathbf{r}) = \left[V^{\text{loc}}(\mathbf{r}) + \frac{1}{2}V_{\text{H}}(\mathbf{r}) + \frac{1}{2}V^{\text{model}}(\mathbf{r})\right]\left[\rho^e(\mathbf{r}) - \rho^{\text{model}}(\mathbf{r})\right]\,, \tag{3.21}$$

where $V^{\text{loc}}(\mathbf{r})$, $V_{\text{H}}(\mathbf{r})$ and $\rho^e(\mathbf{r})$ are already known in real space. Different model potential $V^{\text{model}}(\mathbf{r})$ and model charge density $\rho^{\text{model}}(\mathbf{r})$ can be constructed as long as they obey above constraints. In this work, the model charge density is a polynomial functional with continuous zeroth-, first- and second-order derivatives at 0 and r_c; for $u = r/r_c$,

$$\rho_\mu^{\text{model}}(r) = \begin{cases} \frac{21}{5\pi r_c^3}\left[1 - 10u^3 + 15u^4 - 6u^5\right] & : r < r_c \\ \quad 0 & : r > r_c \end{cases} \tag{3.22}$$

As shown in Figure 3.1, the local charge density varies rapidly with respect to radius r, while the model charge density parametrized in Eqn. (3.22) smoothly decays to zero as increasing r from 0 to the cutoff radius r_c, see Figure 3.2. The corresponding potential (see details in Appendix (D)) is

$$V_\mu^{\text{model}}(r) = \begin{cases} \frac{1}{5r_c}\left[12 - 14u^2 + 28u^5 - 30u^6 + 9u^7\right] & : r < r_c \\ \quad 1/r & : r > r_c \end{cases} \tag{3.23}$$

The model charge density gives faster numerical convergence on a regular spatial grid compared to the rapidly varying local charge density. This model charge density has been tested by calculating the surface energy based on a Si bulk (8 atom) and a Si(111) slab (16 atom). The energy difference between the values calculated from total energy calculation and from energy density integration is about 7 meV. It has also been tested on O atoms, and O_2 molecules for various grid sizes, where we had a convergence problem by using a rapidly-varying local charge density. For a wide range of grid sizes, the total energy calculations converge to a precision of 1 meV, while the difference between results calculated from those two methods is up to 0.4 eV. However, the energy difference can be reduced below 1meV with the smooth model charge.

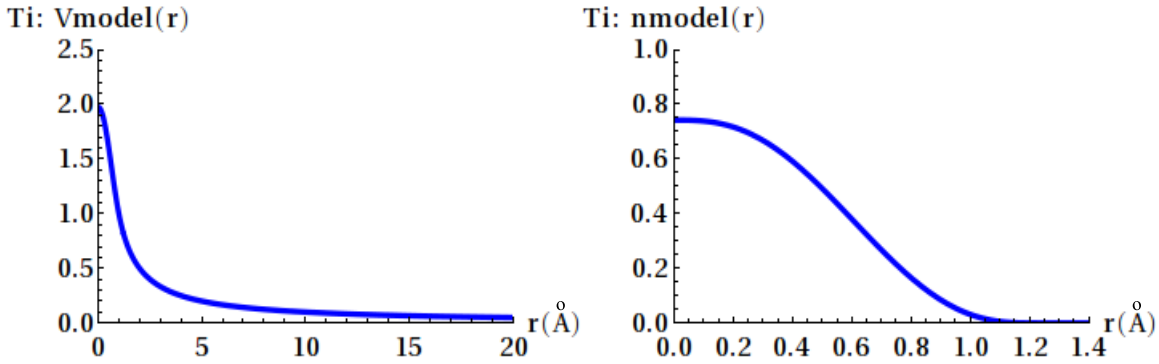

Figure 3.2: Compensating model potential and charge density for PAW Ti with a cutoff radius 1.22 Å. See Appendix (D) for detailed construction. The model potential $V^{\text{model}}(\mathbf{r})$ is a smoothly varying long-ranged potential; the model charge density $\rho^{\text{model}}(\mathbf{r})$ is also a smoothly varying function for the radius r from 0 to the cutoff radius r_c, which integrates to the negative charge density of the local charge.

3.3 Exchange-correlation energy density

The exchange-correlation energy of many-body interacting electrons can be expressed in terms of an exchange-correlation hole which tends to be localized around each electron. In density functional theory it is usually treated as a function of the local density and its

gradients, which is determined by the choice of exchange-correlation functional. For both the local density approximation (LDA) and generalized gradient approximation (GGA), the gauge-independent exchange-correlation energy density is

$$e_{\rm XC}(\mathbf{r}) = \rho^e(\mathbf{r})\,\varepsilon_{\rm XC}\left[\rho^e(\mathbf{r}), |\nabla\rho^e(\mathbf{r})|\right], \tag{3.24}$$

where $\varepsilon_{\rm XC}$ is exchange-correlation energy per electron.

3.4 On-site energy

The last term of the energy density in Eqn. (3.2) is short-ranged. For the PAW method, the on-site energy for each ion is composed of kinetic energy, exchange-correlation energy, and Coulomb energy including electron-electron and electron-ion interactions in the augmentation region, and is

$$E_{\text{on-site}} = (E^1_\mu - \tilde{E}^1_\mu)\delta(\mathbf{r}-\mathbf{R}_\mu), \tag{3.25}$$

with

$$\begin{aligned}
E^1_\mu &= \sum_{ij}\rho_{ij}\langle\phi_i| - \frac{1}{2}\nabla^2|\phi_j\rangle + \overline{E_{\rm XC}[\rho^1+\rho_c]} + \overline{E_{\rm H}[\rho^1]} + \int d\mathbf{r}\, V_{\rm H}[\rho_{Zc}]\,\rho^1\,, \\
\tilde{E}^1_\mu &= \sum_{ij}\rho_{ij}\langle\tilde{\phi}_i| - \frac{1}{2}\nabla^2|\tilde{\phi}_j\rangle + \overline{E_{\rm XC}[\tilde{\rho}^1+\hat{\rho}+\tilde{\rho}_c]} + \overline{E_{\rm H}[\tilde{\rho}^1+\hat{\rho}]} + \int d\mathbf{r}\, V_{\rm H}[\tilde{\rho}_{Zc}](\tilde{\rho}^1+\hat{\rho})\,.
\end{aligned} \tag{3.26}$$

Here the on-site energies E^1_μ and $\tilde{E}^1_\mu$ have been discussed in Section 2.4.3. In practice, one can calculate E^1_μ and $\tilde{E}^1_\mu$ for each ion in terms of radial grids.

For the NCPPs and the USPPs methods, this short-ranged term corresponds to the non-local pseudopotential energy. For the USPPs method, one has

$$E^{\rm nl}_\mu = \sum_{n\mathbf{k}}\int d\mathbf{r}\,\tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r})\left(\sum_{ij}D^{\rm ion}_{ij}|\beta_i\rangle\langle\beta_j|\right)\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})\,. \tag{3.27}$$

For the NCPPs method, the expression is

$$E_{\mu}^{\mathrm{nl}} = \sum_{n\mathbf{k}} \sum_{\ell} \int d\mathbf{r}\, \tilde{\psi}_{n\mathbf{k}}^{*}(\mathbf{r})\, V_{\mu\ell}^{\mathrm{nl}}(|\mathbf{r}-\mathbf{r}_{\mu}|)\, \wp_{\ell}\, \tilde{\psi}_{n\mathbf{k}}(\mathbf{r}). \tag{3.28}$$

3.5 Summary

We summarize all the energy density terms in Table 3.1 for different pseudopotential and the PAW methods. There are 4 terms for the total energy density as defined in Eqn. (3.2). The localized pseudo kinetic energy density generally is written in terms of pseudo–wavefunctions. Both symmetric and antisymmetric forms can be used to evaluate the kinetic energy density. The final results from both forms are identical except terms proportional to the Laplacian of the electron density. This difference is understandable, because those two results are from different gauge choices. After introducing a site-localized charge density, which generates the local pseudopotential, and defining the total classical Coulomb potential as a sum of the Hartree potential and the local pseudopotential, the total classical Coulomb energy density can simply be expressed either in terms of the total Coulomb potential and the total charge density, or in terms of the total electrostatic field squared from the Maxwell expression. The same as the kinetic energy density, the Coulomb energy density is unique up to a gauge transformation. Noticing that a rapidly varying localized charge density brings large numerical errors for the Coulomb energy density, a smooth model charge density has been introduced in the asymmetric form of the Coulomb energy density. The exchange-correlation energy density can be expressed in various functionals, which are all given as a function of the electron density and its gradients. The short-ranged terms in the core regions are treated differently in different methods. We have derived the appropriate forms for the pseudopotential and the all-electron PAW methods, which lead to terms assigned to each atom and not explicitly given as densities.

We have two energy density terms to be integrated which are gauge dependent: the

$e(\mathbf{r}) = t(\mathbf{r}) + e_{\rm CC}(\mathbf{r}) + e_{\rm XC}(\mathbf{r}) + E_{\text{on-site}}\delta(\mathbf{r}-\mathbf{R}_\mu)$.	(3.2)
1. Kinetic energy density	
$t^{(s)}(\mathbf{r}) = \frac{1}{2}\sum_{n\mathbf{k}} f_{n\mathbf{k}} \lvert\nabla\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})\rvert^2$.	
$t^{(a)}(\mathbf{r}) = -\frac{1}{2}\sum_{n\mathbf{k}} f_{n\mathbf{k}}\tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r})\nabla^2\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})$.	(3.3)
$t^{(a)}(\mathbf{r}) - t^{(s)}(\mathbf{r}) = -\frac{1}{4}\nabla^2\tilde{\rho}(\mathbf{r})$.	(3.4)
Construct zero-flux volume Ω_ρ where $\nabla\tilde{\rho}(\mathbf{r})\cdot\hat{n} = 0$	
The bounded volume integral, $T = \int_{\Omega_\rho} t(\mathbf{r})$.	
2. Classical Coulomb energy density	
$e^{\rm Maxwell}_{\rm CC}(\mathbf{r}) = \frac{1}{8\pi}\lvert\nabla V^{\rm tot}(\mathbf{r})\rvert^2$.	(3.13)
$e^{(a)}_{\rm CC}(\mathbf{r}) = \frac{1}{2}V^{\rm tot}(\mathbf{r})\rho^{\rm tot}(\mathbf{r})$.	(3.14)
$= [V^{\rm loc}(\mathbf{r}) + \frac{1}{2}V_{\rm H}(\mathbf{r}) + \frac{1}{2}V^{\rm model}(\mathbf{r})][\rho^e(\mathbf{r}) - \rho^{\rm model}(\mathbf{r})]$.	(3.21)
$e^{(a)}_{\rm CC}(\mathbf{r}) - e^{\rm Maxwell}_{\rm CC}(\mathbf{r}) = -\frac{1}{8\pi}\nabla\cdot(V^{\rm tot}(\mathbf{r})\nabla V^{\rm tot}(\mathbf{r}))$.	(3.15)
Construct zero-flux volume Ω_V where $\nabla V^{\rm tot}(\mathbf{r})\cdot\hat{n} = 0$	
The bounded volume integral, $E_{\rm CC} = \int_{\Omega_V} e_{\rm CC}(\mathbf{r})$.	
3. Exchange-correlation energy density	
$e_{\rm XC}(\mathbf{r}) = \rho^e(\mathbf{r})\,\varepsilon_{\rm XC}(\rho^e(\mathbf{r}))$.	(3.24)
The bounded volume integral, $E_{\rm XC} = \int_{\Omega_\rho} e_{\rm XC}(\mathbf{r})$.	
4. On-site energies	
PAW: $E_{\text{on-site}} = (E^1_\mu - \tilde{E}^1_\mu)$.	(3.25)
USPPs: $E^{\rm nl}_\mu = \sum_{n\mathbf{k}}\int d\mathbf{r}\,\tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r})(\sum_{ij} D^{\rm ion}_{ij}\lvert\beta_i\rangle\langle\beta_j\rvert)\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})$.	(3.27)
NCPPs: $E^{\rm nl}_\mu = \sum_{n\mathbf{k}}\sum_\ell\int d\mathbf{r}\,\tilde{\psi}^*_{n\mathbf{k}}(\mathbf{r})V^{\rm nl}_{\mu\ell}(\lvert\mathbf{r}-\mathbf{R}_\mu\rvert)\wp_\ell\tilde{\psi}_{n\mathbf{k}}(\mathbf{r})$.	(3.28)
The atomic energy: $E = T + E_{\rm CC} + E_{\rm XC} + E_{\text{on-site}}$.	

Table 3.1: Summary of the energy density formulae for PAW, USPPs, and NCPPs methods and the procedure to calculate atomic energy using the energy density method.

kinetic energy density and the classic Coulomb energy density. Defining a gauge independent energy requires integrating these energy densities over volumes to cancel out any gauge dependence. Chetty and Martin [15] used a supercell composed of an integral number of Wigner-Seitz cells; Rapcewicz *et al.* [39] constructed a Voronoi polyhedron for each

comprised atom. However, these volumes are not the *general* solution to removing gauge dependence. For the kinetic energy density, the gauge dependence, Eqn. (3.4), is proportional to the Laplacian of electronic charge density. Hence, we integrate over a volume where the gradients of electron density has zero component along surface normal direction $\hat{n}$, $\nabla\rho(\mathbf{r}) \cdot \hat{n} = 0$—the zero-flux "Bader" volume Ω_ρ. [17] For the classical Coulomb energy density, the gauge dependence, Eqn. (3.15), is proportional to the Laplacian of the potential. Hence, we integrate over a volume where the electrostatic field has zero component along surface normal direction $\hat{n}$, $\nabla V(\mathbf{r}) \cdot \hat{n} = 0$—the zero-flux charge-neutral volume Ω_V. In this work, we construct two different volumes: the Bader volume is used to integrate kinetic energy density and exchange-correlation energy density, and the charge-neutral volume is used to integrate classical Coulomb energy density. Each of these volumes is "atom-centered"—that is, it contains one atom somewhere in the volume—and they each partition space: the union of all volumes is the total supercell volume, and the intersection of any two volumes is zero. We define these volumes on the same regular spatial grid used to represent the charge density and energy density terms. Accurate definition of the volumes and integration uses a weighted integration scheme discussed in Chapter 4 that has quadratic convergence in the grid density. The formulae of the energy density derived in this chapter have many other applications. For example, the well-defined exchange kinetic energy is related to the electron localization function, see Appendix (B).

Chapter 4

Methods and tests for gauge independent integration

In the previous chapter, we derive the energy density formalism which contains two gauge dependent terms: the kinetic energy density and the classical Coulomb energy density. The pseudo kinetic energy density generally can be expressed in symmetric or antisymmetric forms. The difference of these two forms leads to the gauge dependent term of the kinetic energy density, which is proportional to the Laplacian of the electronic charge density. The well-defined kinetic energies can be obtained as integrals over any volume Ω_ρ bounded by a zero-flux surface of the gradient of the electron density, which is the Bader volume [17]. In the middle panel of Figure 4.1 we show the Bader volume of Cl in NaCl crystal. We treat the Coulomb energy in a way analogous to the kinetic energy using the fact that the Coulomb energy density can be expressed in two forms either in terms of the potential and the charge density, or in terms of the electric field squared from the Maxwell expression. Thus it is uniquely defined in a volume bounded by a zero-flux surface of the electrostatic field, i.e., a zero-charge volume Ω_V that is different from the Bader volume. The right panel of Figure 4.1 shows the charge neutral volume of Cl in NaCl crystal. The Cl atom is negative charged in the Bader volume, which is larger than its charge neutral volume. We also should mention that NaCl is an extreme case as an ionic compound. For other cases like the GaAs semiconductor, the Bader volume and the charge neutral volume are not so different, as shown in Figure 5.1.

As mentioned already in Section 1.3 of the introduction, our choice of assigning the Coulomb energy to each atom using the charge neutral volume is different from the Bader's approach, because our definition is more local than the Bader's approach. In our approach,

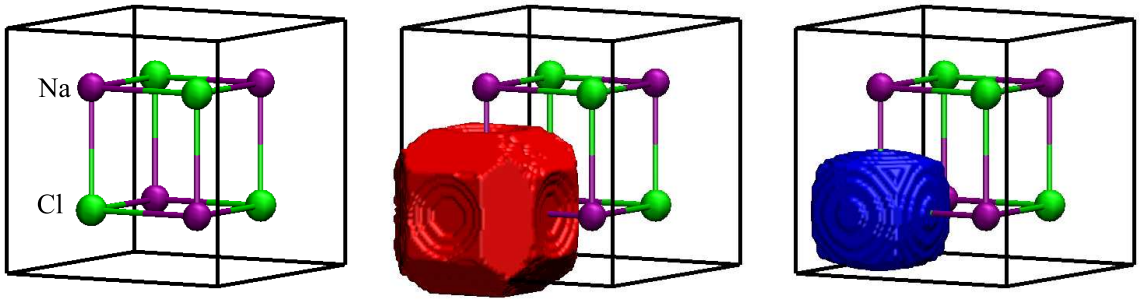

Figure 4.1: Bader volume and charge neutral volume of Cl in NaCl crystal. Left panel shows an eight-atom NaCl cell. The Bader volume in the middle panel has a larger volume than the charge neutral volume in the right panel for the Cl atom.

the Coulomb energy is primarily from interactions of the charges within each region. The interactions among different regions assigned to different atoms are small. Those interactions are at most dipole-dipole, and can even have the lowest order terms to be quadrapole for some high symmetrical cases like NaCl. On the contrary, the Bader's method has interactions of charged units with long range effects, so has no obvious way to assign volumes to individual atoms. Although both methods should give the same total energies, our method provides a more intuitive way to understand different surface and defect energies, and will be explored further in the following.

4.1 Grid-based methods

Each Bader volume, reviewed in Section 1.3, contains a single electron density maximum, and is separated from other volumes by a zero-flux surface of the gradients of the electron density, $\nabla\rho(\mathbf{r}) \cdot \hat{n} = 0$. Here, $\rho(\mathbf{r})$ is the electron density, and $\hat{n}$ is the unit vector perpendicular to the dividing surface at any surface point $\mathbf{r} \in \partial\Omega_\rho$. Each volume Ω_ρ is defined by a set of points where following a trajectory of maximizing ρ reaches the same unique maximum (fixed point). In practical numerical calculations, it is very challenging to have an accurate determination of a zero-flux surface, although different approaches have been proposed to

find a robust and efficient solution.

Different approaches for condensed, periodic systems have relied on analytic expressions of the density [42, 43] or discretizing the charge density trajectories [44, 45, 46, 47, 48]. Early algorithms were based on the electron density calculated from analytical wavefunctions of small molecules, and integration along the gradient paths. Most current developments are based on a grid of electron density, which is important for DFT calculation and also applicable to analytical density function of small molecules. One octal tree algorithm [44] uses a recursive cube subdivision to find the atomic basins robustly, but practically is not applicable to complicated topologies due to huge computational cost. The "elastic sheet" method [45] defines a series of fictitious particles which gives a discrete representation of zero-flux surface. Particles are relaxed according to the gradients of electron density and interparticle forces. This method will not work for complex surfaces with sharp cusps or points.

Recently, Henkelman *et al.* developed an on-grid method [46] to divide an electron density grid into Bader volumes. This method can be applied to the DFT calculations of large molecules or materials. They discretize the trajectory to lie on the grid, ending at the local maximal point of the electron density. The points along each trajectory are assigned to the atom closest to the end point. Although this method is robust, and scales linearly with the grid size, it introduces a lattice bias caused by the fact that ascent trajectories are constrained to the grid points. The near-grid method [47] improves this by accumulating a correction vector—the difference between the discretized trajectory and the true trajectory—at each step. When the correction vector is sufficiently large, the discrete trajectory is corrected to a neighboring grid point. This method corrects the lattice bias, and also scales linearly with respect to the size of grids. However, both grid trajectory methods require iteration to self-consistency in volume assignments. Also, the integration error scales linearly with the grid spacing, so very fine grids are required in numerical calculations to provide the correct Bader volume, reducing its applicability for accurate calculations in a large system.

Lastly, a new algorithm uses a "divide and conquer" adaptive approach with tetrahedra; tetrahedra are continuously divided at the boundaries of Bader volumes, with the weight of each tetrahedra given by the number of vertices that belong to each volume [48]. Such an approach retains linear scaling with the grid spacing, but requires mesh refinement near boundaries to deal with the linear convergence of the error with the grid spacing.

Here we show an example of misassignment of the grid points to basins from the near-grid method, where misassignment occurs for the grid points close to the dividing surfaces with the gradients of charge density almost parallel to the surfaces. In this example, a three dimensional model charge density is constructed from three Gaussian functions in simple cubic unit cell, $\rho(\mathbf{r}) = \sum_{i=1,3} e^{(-\mathbf{r}-\mathbf{r}_i)^2/W^2}$. The $\mathbf{r}_i$ are $(0.25N, 0.25N, 0.4N)$, $(0.5N, 0.5N, 0.5N)$, and $(0.75N, 0.75N, 0.4N)$, with grid size $N^3 = 60^3$ and width $W = N/10$. Figure 4.2 shows the charge density distribution on $(1\bar{1}0)$ plane. Due to the symmetry of charge density distribution, the true dividing surfaces along charge density saddle points are known analytically and shown as two black lines on $(1\bar{1}0)$ plane. As shown in the top panel of Figure 4.3, the grid points marked by orange circles are assigned to the wrong basins by the near-grid method, different from the partition of spatial points by true dividing surfaces. The gradients of charge density shown in arrows for these misassigned surface points have small normal components, and we believe this is the cause of the misassignment.

Figure 4.4 illustrates the partition of the real space into two Bader volumes A and B by a zero-flux dividing surface. Arrows denote the gradients of charge density (a). The component of $\nabla\rho$ along the surface normal $\hat{n}$ is zero for any point on the surface ∂A or ∂B. A grid-based partition algorithm (b), such as the near-grid method, divides space into volume surrounding around each grid point, and assigns each grid volume to a particular Bader volume. Even though grid points may be assigned to Bader volumes correctly, the density integration based on the grid-based partition would bring in numerical integration error that scales linearly with the grid spacing. Introducing a "weight" integrand (c) representing the fraction of grid volume that belongs to a particular Bader volume smooths out the grid-based partition,

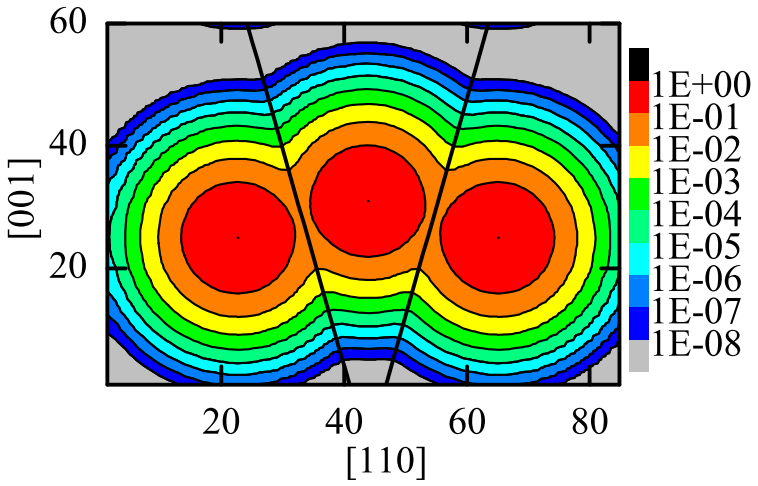

Figure 4.2: Charge density distribution constructed from three Gaussian functions. The grid size is $N = 60$, and the charge density is shown on $(1\bar{1}0)$ plane; this is coplanar with the centers of the three Gaussian functions. The true dividing surfaces are indicated by two black lines due to symmetry.

and improves the integration accuracy and scales quadratically with the grid spacing. The atomic contribution is neither 1 nor 0 at the dividing surfaces, but fractional. In Figure 4.4, red represents a weight of 1 to atom A for grid points closer to atom A, and transitions to white for a weight of 0 for grid points away from atom A.

In this work, we present a new, efficient, accurate method to carry out the integration of energy density or other charge density relevant functional. The weight method is applied to calculating the integration of energy density bounded by zero-flux surface of the gradients of electron density or the gradients of Coulomb potential, generally is applied on constructing zero-flux surface of the gradients of an arbitrary functional. The weight functionals represent volume fractions of the cell of each grid point assigned to different atoms. The algorithm is robust, efficient with $O(N)$ computing time, and more accurate than other existing grid-based algorithm. This method depends upon the formulation of flow across the dividing surfaces between cells that can be applied to uniform or non-uniform grids.

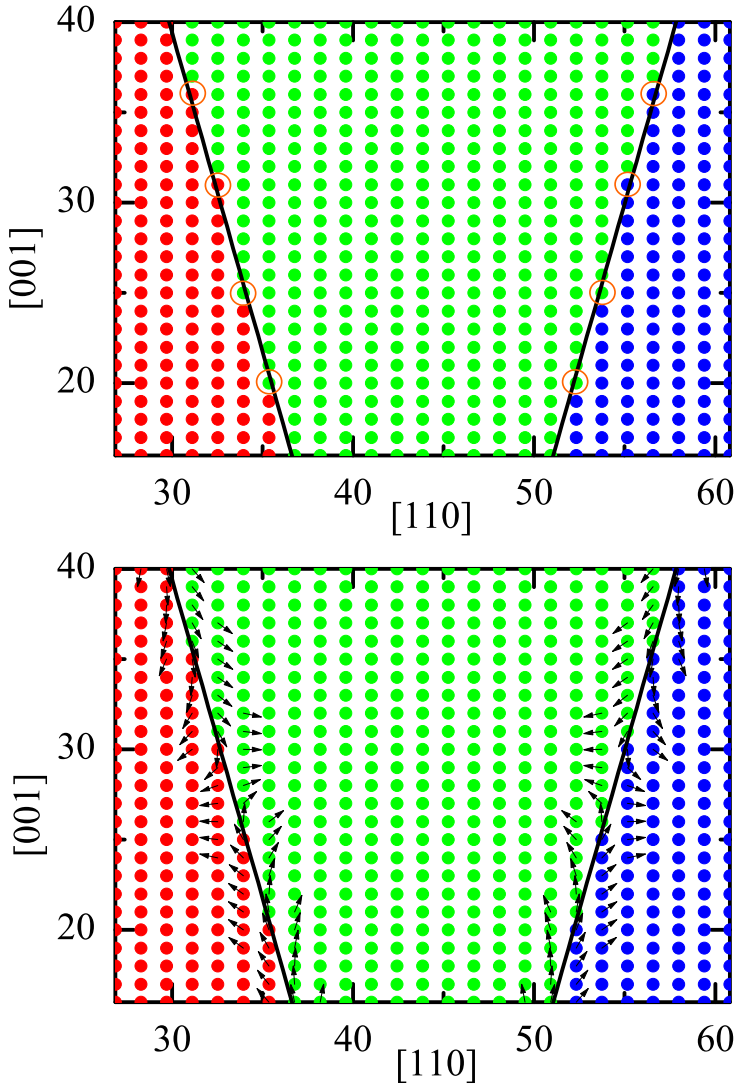

Figure 4.3: Basin identification with the near-grid and the weight methods for a three-dimensional Gaussian functional charge density. 2D plot is the $(1\bar{1}0)$ plane of a simple cubic cell. The basin assignment of grid points is given: red dots for ion I, green dots for ion II, and blue dots for ion III; for the weight method, a single color is assigned to the maximum weight at each point. Basin assignment from the near-grid method is given in the top panel; basins with maximal weight on every grid points from the weight method are indicated in the bottom panel. Orange circles in the top panel indicate the grid points misassigned by the near-grid method and corrected by the weight method. Arrows in the bottom panel denote the directions of the gradients of charge density, which can be used to verify the correctness of basin assignment.

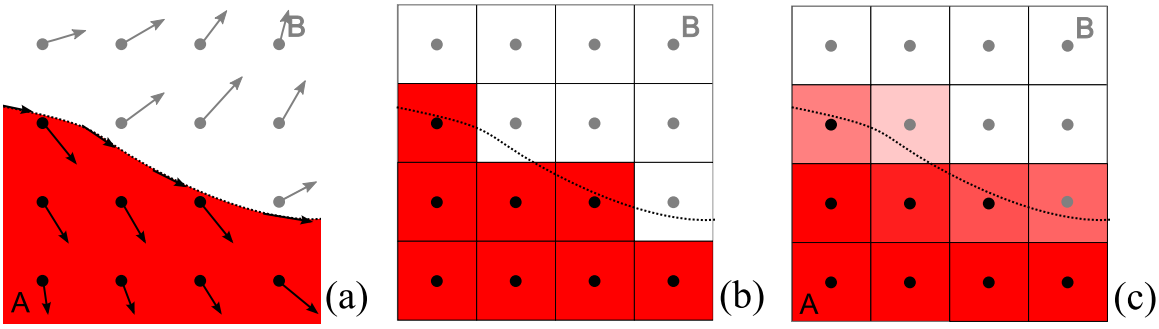

Figure 4.4: Schematic illustration of the zero-flux surface, the near-grid algorithm, and weighted integration. The zero-flux dividing surface separate volumes A and B, where arrows denote charge density gradients (a). The normal component is zero for any point on the surface $\nabla\rho \cdot \hat{n} = 0$. The near-grid method [47] gives grid-based partition (b), however energy density integration based on this grid-based partition would cause integration error due to finite grid sizes. A weight function (c) representing volume fractions of the cell of each grid point is introduced to reduce the error due to a finite grid.

4.2 The weight method

The Bader partitioning of space defines volumes by the endpoint of a trajectory following the gradient flow of the charge density, $\nabla\rho$. We assume that ρ has continuous first and second derivatives throughout all space of interest, and has a set of discrete local maxima (fixed points) $\mathbf{x}_1$, $\mathbf{x}_2$, etc., where $\nabla\rho = 0$ and the matrix $\nabla\nabla\rho$ is negative-definite. The basin of attraction A_n, of a fixed point $\mathbf{x}_n$ is the set of points which flow to the fixed point $\mathbf{x}_n$ along the charge density gradient. That is, for any point $\mathbf{r}$, we can integrate the trajectory given by $\dot{\mathbf{x}}(t) = \nabla\rho(\mathbf{x})$, with the initial condition $\mathbf{x}(0) = \mathbf{r}$, to find $\lim_{t\to\infty} \mathbf{x}(t)$. Each trajectory will end at fixed point $\mathbf{x}_n$, and except for a set of points with zero volume in space, the extremum is a local maximum; the basin of attraction A_n are all points $\mathbf{r}$ whose trajectory $\lim_{t\to\infty} \mathbf{x}(t)$ ends at $\mathbf{x}_n$. Note also that if point $\mathbf{r}_0 \in A$, and the trajectory starting from $\mathbf{r}_1$ reaches $\mathbf{r}_0$ in a finite time t, then $\mathbf{r}_1 \in A$. This set defines a partitioning of space, where $A_n \cap A_m = \emptyset$ when $n \neq m$ and $\cup_n A_n = \Omega$. Finally, each basin A_n is such that wherever the normal $\hat{n}$ to the bounding surface ∂A_n is well-defined, $\hat{n} \cdot \nabla\rho = 0$. If ρ is the charge density, then A_n are the Bader volumes; but this definition is applicable to any sufficiently

smooth function with a discrete set of local maxima. As the definition of the basins A_n derives from trajectories, it is not possible in general to determine if two neighboring points $\mathbf{r}$ and $\mathbf{r}'$ belong to the same or different basins based only on local information.

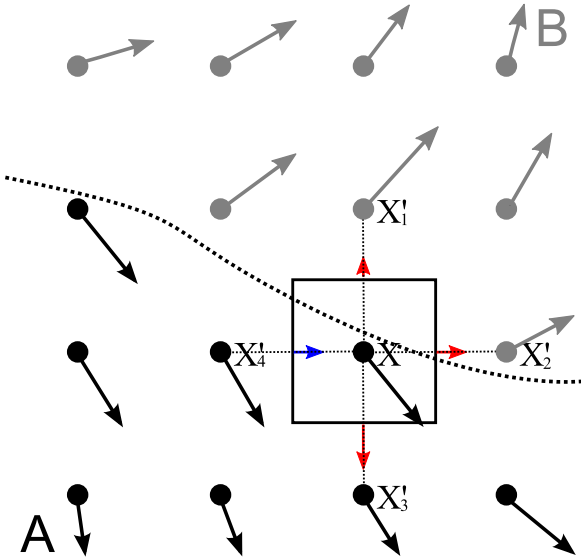

Figure 4.5: Schematic illustration of the weight method. The zero-flux dividing surface separate volumes A and B, where black and grey arrows denote charge density gradients on grid points. The volume of the cell of a grid point flows to its neighbors with larger charge density magnitude. Flowing flux is shown as directional map, either flowing from X to X' as red arrows or flowing from X' to X as blue arrows.

Figure 4.5 shows the reformulation for an approximate fractional partitioning of real-valued function evaluated at a set of discrete points, X. The grid points X partition space into Voronoi polyhedra [49] V_X covering each grid point X, where a point in space $\mathbf{r}$ belongs to the volume V_X if X is the closest point in Cartesian space to $\mathbf{r}$. Each polyhedra is defined by the nearest neighboring points X' that are a distance $\ell_{X\to X'}$ away; the Voronoi polyhedron at X has facets $\partial V_{X\to X'}$ with normal $\hat{n}_{X\to X'}$ pointing from X to X' and area $a_{X\to X'}$. Moreover, the facet is at the midpoint between X and X'. Our goal is to define for each grid point X, a "weight" $w^A(X)$ between 0 and 1 such that $\sum_A w^A(X) = 1$ for all X,

and the discrete approximation to the integral over the basin A

$$\int_A d\mathbf{r}\, f(\mathbf{r}) \approx \sum_X V_X w^A(X) f(X) \tag{4.1}$$

converges quadratically in the grid spacing for smooth functions $f(\mathbf{r})$. The weight, in this case, is the fraction of points in V_X whose trajectory ends in the basin A. Note that if the points X form a regular periodic grid, the Voronoi volumes, facet areas, and neighbor distances need only be computed for the Wigner-Seitz cell around a grid point.

To transition from the continuum definition of spatial partitioning to our Voronoi partitioned definition, we introduce the continuum probability density for our trajectories, $P(\mathbf{r}, t)$. From the trajectory equation, the probability flux at any point and time is $\mathbf{j}(\mathbf{r}, t) = P(\mathbf{r}, t)\nabla\rho(\mathbf{r})$. Then, the probability distribution evolves in time according to a continuity equation

$$\frac{\partial P(\mathbf{r}, t)}{\partial t} + \nabla \cdot (P(\mathbf{r}, t)\nabla\rho(\mathbf{r})) = 0\,. \tag{4.2}$$

This equation represents the combined evolution of a distribution of points in space; we use it to determine how the points in V_X distribute to neighboring volumes $V_{X'}$. Define the volume probability

$$P_X(t) = V_X^{-1} \int_{V_X} d\mathbf{r}\, P(\mathbf{r}, t); \tag{4.3}$$

then the evolution from the initial condition

$$P(\mathbf{r}, 0) = \begin{cases} 1 & : \mathbf{r} \in V_X \\ 0 & : \mathbf{r} \notin V_X \end{cases} \tag{4.4}$$

is given by

$$\begin{aligned}
\frac{dP_X(t)}{dt} &= -V_X^{-1}\int_{V_X} d\mathbf{r}\nabla\cdot(P(\mathbf{r},t)\nabla\rho(\mathbf{r})) \\
&= -V_X^{-1}\sum_{X'}\int_{\partial V_{X\to X'}} P(\mathbf{r},t)\nabla\rho\cdot\hat{n}_{X\to X'}dS \\
&\approx -V_X^{-1}P_X(t)\sum_{X'}\int_{\partial V_{X\to X'}} \nabla\rho\cdot\hat{n}_{X\to X'}dS \\
&\approx -P_X(t)\sum_{X'}\frac{a_{X\to X'}}{V_X}\cdot\frac{R(\rho_{X'}-\rho_X)}{\ell_{X\to X'}} \\
&\equiv -P_X(t)\sum_{X'}\tau_{X\to X'}\,,
\end{aligned} \tag{4.5}$$

where $R(u) = u\theta(u)$ is the ramp function, so that $\tau_{X\to X'} \geq 0$ and is zero when $\rho_{X'} < \rho_X$; this is a consequence of our initial conditions where $P(\mathbf{r},0)$ is only nonzero in the interior of V_X. The first approximation ignores spatial variation of $P(\mathbf{r},t)$ through the volume V_X (an error linear in the grid spacing), and the second approximation ignores spatial variation of $\nabla\rho$ along a facet $\partial V_{X\to X'}$, and approximates the gradient at the midpoint between X and X' with the finite difference value (also with an error that is linear in the grid spacing). The solution to Eqn. (4.5) is $P_X(t) = \exp(-t\sum_{X'}\tau_{X\to X'})$. For that solution, the time-integrated flux of probability from V_X to $V_{X'}$ through the facet $\partial V_{X\to X'}$ is

$$\begin{aligned}
J_{X\to X'} &= \int_0^\infty dt\int_{\partial V_{X\to X'}} P(\mathbf{r},t)\nabla\rho\cdot\hat{n}_{X\to X'}dS \\
&\approx \int_0^\infty dtP_X(t)\int_{\partial V_{X\to X'}} \nabla\rho\cdot\hat{n}_{X\to X'}dS \\
&\approx \int_0^\infty dtP_X(t)\tau_{X\to X'} \\
&= \frac{a_{X\to X'}\ell_{X\to X'}^{-1}R(\rho_{X'}-\rho_X)}{\sum_{X'}a_{X\to X'}\ell_{X\to X'}^{-1}R(\rho_{X'}-\rho_X)}\,,
\end{aligned} \tag{4.6}$$

where we have used the same approximations as above. This flux defines the total fraction of points inside V_X that transition to volume $V_{X'}$ through $\partial V_{X\to X'}$, shown as red arrows in Figure 4.5. Note that $\sum_{X'} J_{X\to X'} = 1$, unless X is a local (discrete) maxima, where

$\rho_X > \rho_{X'}$ for all neighbors X'. Finally, as the weight $w^A(X)$ represents the volume fraction of points in volume V_X whose trajectory ends inside basin A, then

$$w^A(X) = \sum_{X'} J_{X\to X'} w^A(X') \,. \tag{4.7}$$

Note that if for all X' where $\rho(X') > \rho(X)$, $\sum_A w^A(X') = 1$, then as $\sum_{X'} J_{X\to X'} = 1$, Eqn. (4.7) guarantees that $\sum_A w^A(X) = 1$.

Forward substitution solves Eqn. (4.7) after the grid points are sorted from highest to lowest density $\rho(X)$. Sequentially, each point X is either

1. A local maxima: $\rho(X) > \rho(X')$ for all neighbors X'. This grid point corresponds to a new basin A, and we assign $w^A(X) = 1$.

2. An interior point: for all X' where $\rho(X') > \rho(X)$, the weights have been assigned and $w^A(X') = 1$ for the *same* basin A. Then Eqn. (4.7) assigns X to basin A as well: $w^A(X) = 1$.

3. A boundary point; with weights between 0 and 1 for multiple basins assigned by Eqn. (4.7).

Then $w^A(X)$ is known from $w^A(X')$ where $\rho(X') > \rho(X)$ for each basin A (as $J_{X\to X'} \neq 0$ only if $\rho(X') > \rho(X)$). Note also that the weight for a particular basin A_n is assigned *without* reference to any other basin A_m; once the set of time-integrated fluxes $J_{X\to X'}$ are known and the densities sorted in descending order, the solution for each basin is straightforward, and Eqn. (4.7) is only needed on the boundary points.

This algorithm solves several issues with the near-grid method. It requires no self-consistency, which improves the computational scaling. Moreover, the introduction of smooth functions that define the volume fraction of points in each basin produces less error and faster convergence with additional grid points. The algorithm is also readily applicable to non-uniform grids, such as an adaptive meshing scheme—it only requires computation of the

Voronoi volumes and facets for the grid points. In one dimension, Eqn. (4.1) has quadratic convergence in the grid spacing; we now demonstrate the quadratic convergence and improved integration accuracy for three dimensional problems.

4.3 Test cases

One determination of the accuracy of Bader volume integration is the vanishing of the volume integration of the Laplacian of charge density $\nabla^2\rho(\mathbf{r})$. The non-zero value of the Laplacian of charge density integration within each Bader volume is our atomic integration error, and can be used as an estimate of the error of the integration of the kinetic energy. We construct the zero-flux surface of the gradients of charge density and evaluate the integration error with both the weight- and near-grid methods for several cases. First, we consider an analytic charge density with known boundaries in an orthogonal and a non-orthogonal cell. Next, we calculate real systems: an ionic compound, and a semiconductor. We also evaluate the Bader charge of Na in NaCl crystal by integrating the charge density within Bader volume, and compare the convergence with the near-grid method.

4.3.1 One dimensional charge density from two Gaussian functions

One-dimensional periodic functional can be a fundamental example to illustrate the convergence with respect to the grid sizes. In one dimension, the volume of Voronoi cell of each grid point i includes the region from $i - \frac{1}{2}$ to $i + \frac{1}{2}$. $J_{i\rightarrow i+di}$ gives the probability of volume flowing from V_i to its neighbor V_{i+di}, where di takes value $\{-1, 1\}$.

In this example, one-dimensional charge density is constructed from two Gaussian functions $\rho(x) = e^{(-x-x_1)^2/W^2} + 2e^{(-x-x_2)^2/W^2}$ with x_1, x_2 located at $0.2N, 0.6N$. N is the grid size and width W is chosen as $0.15N$. The top right inset of Figure 4.6 shows the charge density distribution on grid size $N = 100$. The gradients of charge density have two zero

points in addition to two Gaussian function centers. These two zero gradient points t_1, t_2 are used to partition two Gaussian functions. Left inset is an example of defining dividing point between grid points i and $i+1$. The gradient of charge density is less than 0 at i, and is larger than 0 at $i+1$. For $t < \frac{1}{2}$ in this example, $\frac{1}{2} - t$ means the volume fraction of Voronoi cell of grid point i flowing to grid point $i+1$. In our algorithm, it can be evaluated by $\frac{J_{i\to i+1}}{J_{i\to i+1}+J_{i\to i-1}}$.

We perform convergence test on grid-based integration and the weight method in terms of various grid sizes ranging from 50 to 10000 grid points. The bottom panel of Figure 4.6 shows the atomic integration error in terms of various grid sizes. The weight method reports smaller integration error than grid-based integration systematically. Fitting data to a non-linear functional, one obtains that grid based integration converges as $2.2 \times 10^{-4}\, N^{-0.93}$, while the weight method converges as $1.7 \times 10^{-3}\, N^{-2.0}$. The weight method is more efficient and more accurate than grid-based integration.

4.3.2 Three dimensional charge density from three Gaussian functions in a simple cubic cell

In this example, a three dimensional model charge density is constructed from three Gaussian functions in a simple cubic unit cell, as shown in Figure 4.2. The model charge density is given in an analytical functional, and thus the Laplacian of charge density can be calculated as an analytical functional, from Fourier transform or finite difference method. A set of charge density grids ranging from 20^3, 40^3, 60^3, 80^3, 100^3 points to 200^3 points are calculated. Figure 4.7 shows the integration of the Laplacian of charge density in the unit cell vanish with numerical error and round-off error of 10^{-17} to 10^{-15} for various grid sizes. The bottom panel of Figure 4.7 compares the atomic integration error from the Bader partition using the near-grid method and the weight method. The weight method reports maximal atomic integration error four order of magnitude lower than the near-grid method systematically .

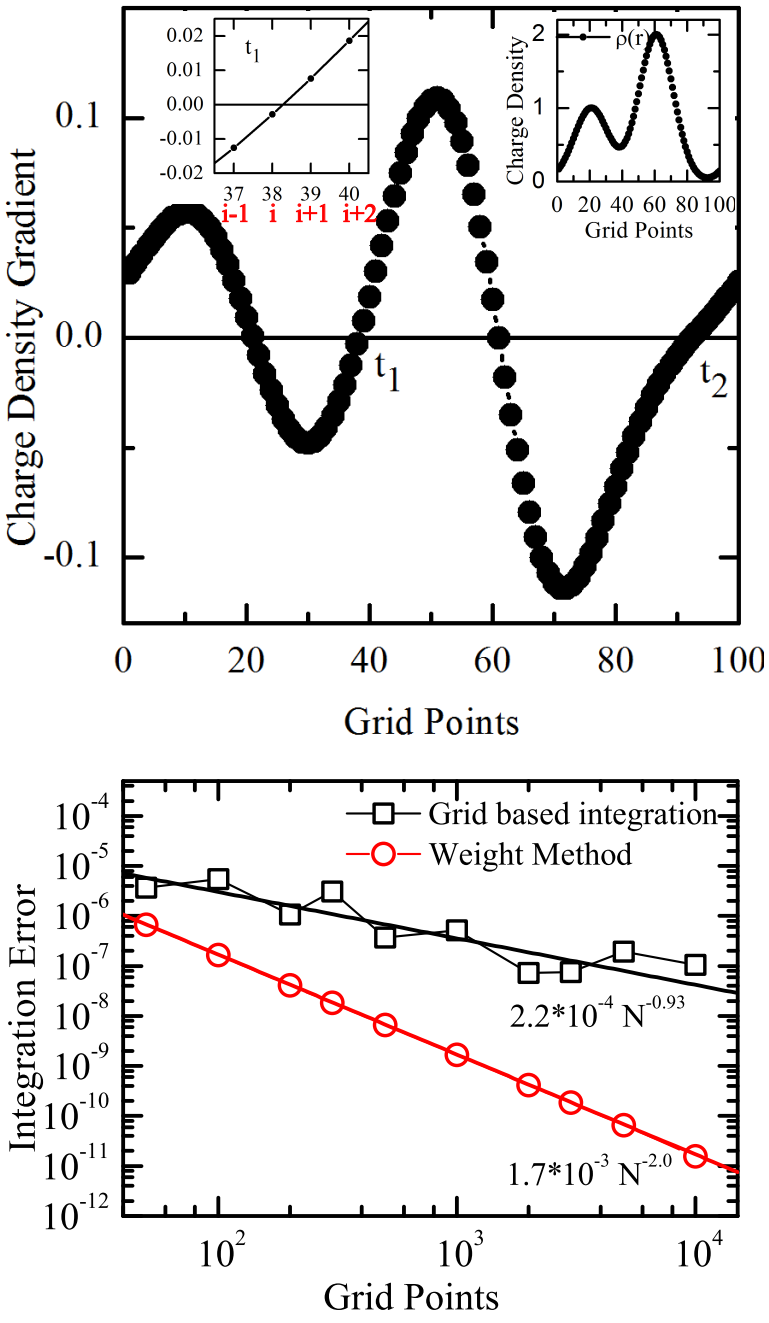

Figure 4.6: Improvement of integration accuracy using the weight method on one-dimensional Gaussian functional charge density. One-dimensional charge density (right inset of the top panel) is plotted on grid size $N = 100$. The two zero gradient points t_1, t_2 are used to partition two Gaussian functions. The bottom panel displays the behavior of integration error on grid-based integration and the weight method with various grid sizes ranging from 50 to 10000 grid points. The weight method converges faster than grid based integration.

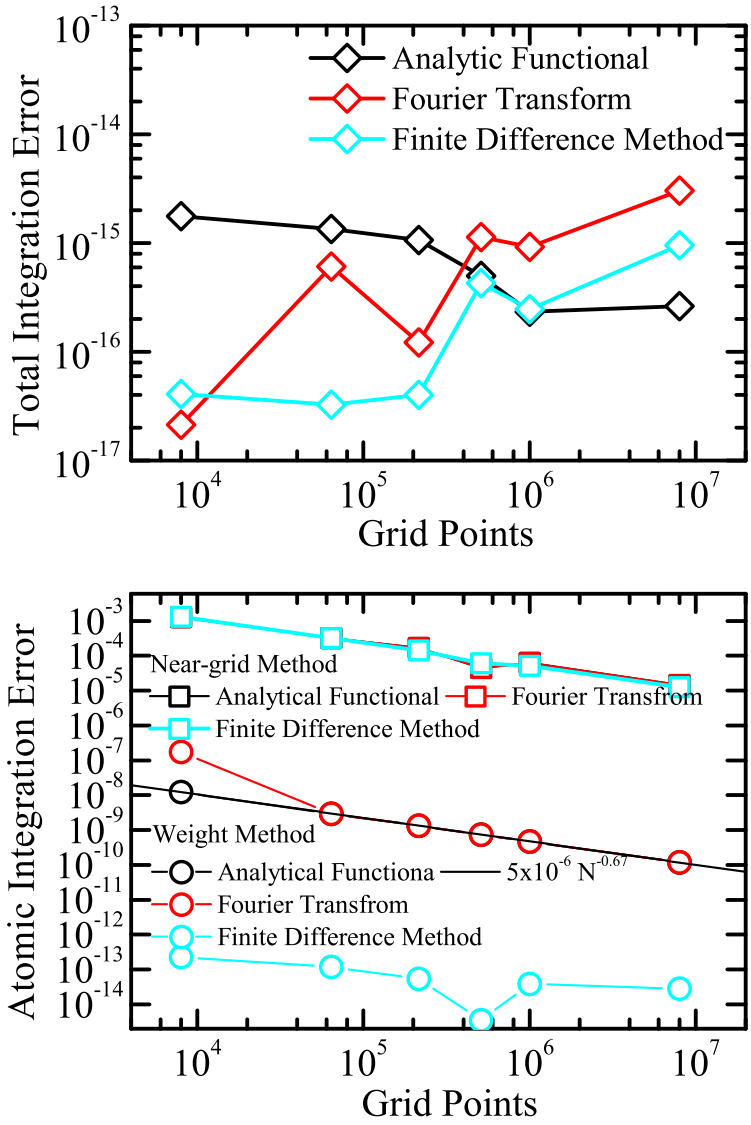

Figure 4.7: Comparison of the near-grid method and the weight method for integration error of Gaussian functional charge density in a simple cubic cell. The Laplacian of charge density are evaluated analytically, by Fourier transform and by finite difference method. Total integration error in the simple cubic unit cell is shown in the top panel. The maximal atomic volume integrations of the Laplacian of charge density within Bader volumes using the near-grid method and the weight method are denoted by squares and circles in the bottom panel. We calculate charge density grids ranging from 20^3 points to 200^3 points. The weight method reports maximal atomic integration error four orders of magnitude lower than near-grid method. The analytical functional and Fourier transform exhibit very close maximal atomic integration errors, while error cancellation exists in the finite difference method.

4.3.3 Three dimensional charge density from three Gaussian functions in a FCC cell

The weight method can be generalized to studying the non-orthogonal charge grids. In this test, we give an example of charge density distributed in a FCC unit cell shown in Figure 4.8. Same as in the last test, the three dimensional model charge density is constructed from three Gaussian functions, $\rho(\mathbf{r}) = \sum_{i=1,3} e^{(-\mathbf{r}-\mathbf{r}_i)^2/W^2}$. The $\mathbf{r}_i$ are located at $(0.25N, 0.25N, 0.4N)$; $(0.5N, 0.5N, 0.5N)$; $(0.75N, 0.75N, 0.4N)$ where N^3 is the number of grid points in the FCC unit cell and width $W = N/10$. We vary N from 20 to 100. Different from simple cubic lattice, the Voronoi cell of FCC lattice has 12 neighbors, where all facets have the same area. The atomic weights on every grid represents the fraction of Voronoi volume of that grid point flowing to specific atom through its neighbors. By calculating on a set of grid sizes, one obtains the maximal atomic integration errors from the near-grid method and the weight method.

Figure 4.8 shows a reduction in error of three orders of magnitude from the near-grid method. Fitting data to a non-linear function $y = a\,N^{-r}$ gives a convergence rate of 0.71 for the weight method, and 0.45 for the near-grid method. The exponent of 0.71 is close to the 2/3 expected for quadratic convergence, and 0.45 is close to the 1/3 expected for linear convergence. The weight method has both better absolute error and converges faster than the near-grid method; in addition, there is no crossover point at large grid spacing where near-grid has smaller errors.

4.3.4 TiO_2 bulk

For a real charge density, we perform DFT calculations on TiO_2 bulk by use of the projector augmented wave (PAW) [25] method, the GGA with PBE functional [22] for the exchange-correlation energy. Atomic configurations for Ti and O are [Ne]$3s^2 3p^6 4s^2 3d^2$ with cutoff radius 1.22 Å, and [He]$2s^2 2p^4$ with cutoff radius 0.58 Å, respectively. We use a plane-wave

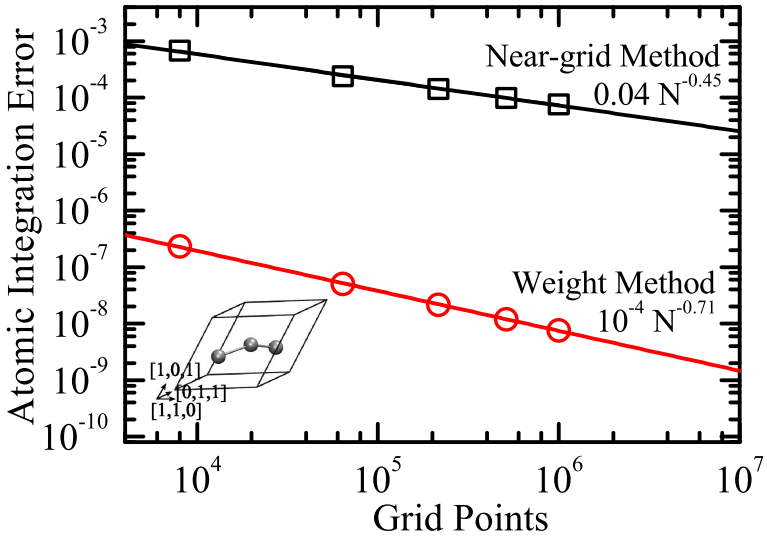

Figure 4.8: Comparison of the near-grid method and the weight method for integration error of Gaussian functional charge density in a FCC cell. The maximal atomic volume integrations of the Laplacian of charge density within Bader volumes using the near-grid method and the weight method are denoted by squares and circles, respectively. We calculate charge density grids ranging from 20^3 points to 100^3. Our algorithm gives atomic integration errors three orders of magnitude lower than the near-grid method, and converges faster than the near-grid method.

basis set with cut-off energy of 900eV. The tetragonal unit cell of rutile TiO_2 (see Figure 4.9) contains two Ti atoms and four O atoms. Monkhorst-Pack k-point method with $4 \times 4 \times 6$ k-points for six-atom cell is used for Brillouin-zone integration with a Gaussian smearing of 0.1eV for electronic occupancies. Theoretically optimized lattice constant are $a = 4.649\,\text{Å}$, $c = 2.970\,\text{Å}$, $u = 0.305$ agreeing with experimental lattice constants of $a = 4.584\,\text{Å}$, $c = 2.953\,\text{Å}$, $u = 0.305$ [6]. A set of charge density grids ranging from $45 \times 45 \times 30$, $60 \times 60 \times 40$, $75 \times 75 \times 50$, $90 \times 90 \times 60$, $120 \times 120 \times 80$ points to $150 \times 150 \times 100$ are calculated. For the energy cutoff of 900 eV, a grid of $45 \times 45 \times 30$ is required to eliminate wrap-around errors, and is the minimum size used by an accurate VASP calculation.

Figure 4.9 shows maximal atomic integration errors as a function of grid sizes. The weight method gives maximal atomic integration error one order of magnitude lower than the near-grid method systematically. The atomic integration error larger than 1.0 eV on the

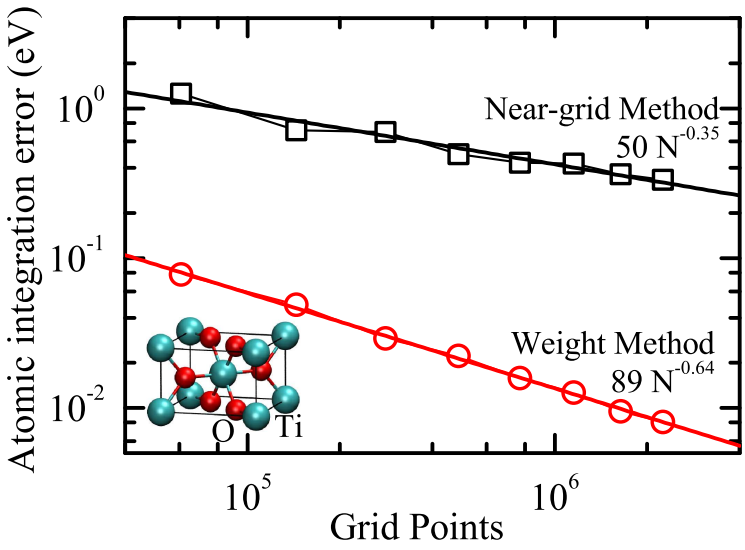

Figure 4.9: Maximal atomic integration error on rutile TiO_2 with respect to the charge density grids. A set of charger density grids ranging from $45\times45\times30$ points to $150\times150\times100$ points are calculated. The weight method reports maximal atomic integration error at least one order of magnitude smaller than the near-grid method. The weight method is practically useful for calculation of small grid size.

minimal grid size $45 \times 45 \times 30$ from the near-grid method is unacceptably large. Again, the convergence rate of the error goes as $\sim 2/3$ for the weight method—corresponding to quadratic convergence—and $\sim 1/3$ for the near-grid method—corresponding to linear convergence. Both the improved error and faster convergence allows for more accurate density integration with fewer grid points than near-grid.

4.3.5 NaCl crystal

In this example, we evaluate the Bader charge (valence electron density integration within Bader volume) of Na atom in NaCl crystal by integrating the charge density within Bader volume, and compare the value with the near-grid method. We perform DFT calculations by use of the PAW method, the GGA with PW91 functional [21] for the exchange-correlation energy. Atomic configurations for Na and Cl are $[He]2s^22p^63s^1$ with cutoff radius 0.77 Å, and $[Ne]3s^23p^5$ with cutoff radius 1.00 Å, respectively. A plane-wave basis set with cut-off energy

of 500 eV is applied. The NaCl unit cell contains 4 Na atoms and 4 Cl atoms. Monkhorst-Pack k-point method with $3 \times 3 \times 3$ k-points for eight-atom cell is used for Brillouin-zone integration with a Gaussian smearing of 0.2 eV for electronic occupancies. The optimized lattice constant of 5.67 Å agrees with the experimental lattice constant of 5.64 Å. A set of charge density grids of 60^3, 80^3, 100^3, 120^3, to 180^3 points are calculated.

Figure 4.10 shows the maximal atomic integration error as a function of various grid sizes. The weight method again shows maximal atomic integration error at least one order of magnitude lower than the near-grid method systematically. The scaling of the error goes as the $\sim 2/3$ power for the weight method, showing continued quadratic convergence, while the near-grid method error scales as the $\sim 1/3$ power, which is linear convergence.

Figure 4.11 shows that Bader charge of Na atom evaluated on various charge grids. The weight method computes a Bader charge of Na atom slightly larger than the near-grid method. Fitting the data to $\rho = \rho_0 + \frac{C}{N_{grid}^{\alpha}}$, we find converged Bader charge values of 0.878 e, 0.881 e, for the near-grid method and the weight method, respectively. We believe this is due to a systematic misassignment for the near-grid method, as shown for the Gaussian charge density case. This suggests that the misassignment may not be improved by increasing the density of grid points in the near-grid method. This suggests that a "divide and conquer" approach using continually refined grids can face potential difficulty. For 60^3 grid points, the near-grid method underestimates the Bader charge by 0.01 e, while the weight method underestimates it by 0.005 e, again showing faster convergence.

4.4 Computational effort

The weight method is computationally efficient, requiring overall effort that scales linearly with the number of grid points. The total computer time is comprised of two primary tasks: the sorting of charge density costs $O(N \log N)$ with N grid points, and the atomic weight evaluation on the sorted grid points beginning from grid point with maximum density

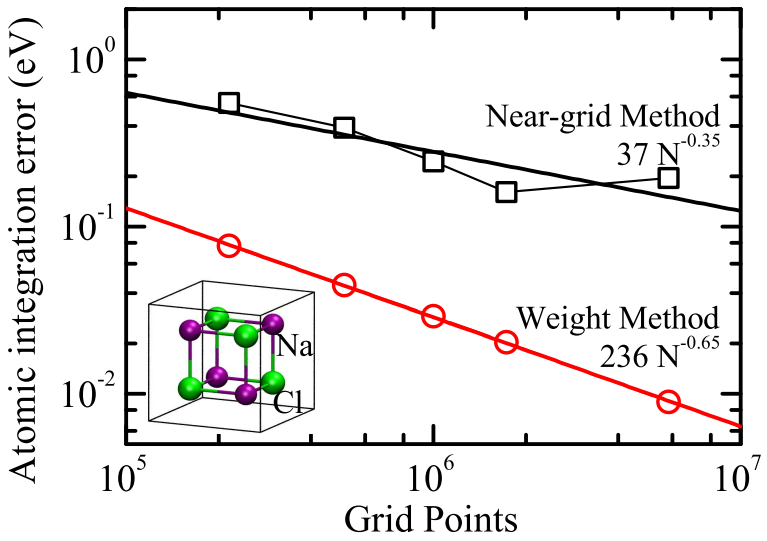

Figure 4.10: Comparison of the near-grid method and the weight method for maximal atomic integration error of NaCl crystal. A set of charge density grids ranging from 60^3 points to 180^3 points are calculated. Comparing to the near-grid method, the weight method reduce the integration error remarkably

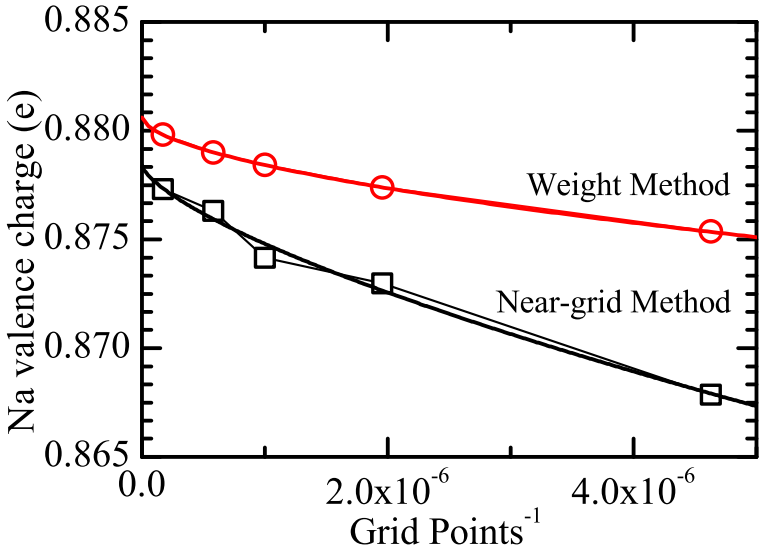

Figure 4.11: Comparison of the near-grid method and the weight method on convergence of Bader charges of Na in NaCl crystal. The Bader charge of Na is calculated for a set of density grids ranging from 60^3 points to 180^3 points. Both methods give monotonic, and smooth convergence.

requires at most $N \times N_{\text{atom}}$ computer time. The computational effort is smaller than that, as only the surface grid points which have fractional atomic weights require N_{atom} calculations, while each interior grid point require only one calculation. Generally, the number of surface grid points is a small fraction of the number of total grid points, and scales as $N^{2/3}$. For example, the ratio of the number of surface grid points to the number of total grid points is 14% in NaCl crystal with total grid sizes 60^3.

In a calculation with charge density grid sizes approaching 10^7–10^8 grid points and up to hundreds of atoms in large supercells, we find our algorithm is not only more accurate, but more efficient than the near-grid method. Both methods scale linearly with the number of grid points. Figure 4.12 shows the linear scaling of computer time required to analyze the charge density grid for an eight-atom NaCl with the number of grid points. The improved efficiency of our algorithm appears to originate from the lack of a self-consistent refinement of basin assignment. Comparing to the near-grid method, which needs refinement integration, our weight method has small prefactor, although both are linearly scaled.

4.5 Summary

In this chapter, we develop a weight method to integrate functions defined on a discrete grid over basins of attraction (such as Bader volumes) in an efficient and accurate manner. The weight method works with the density on a discrete grid and assigns volume fractions of the Voronoi cell of each grid point to surrounding basins. Starting from the local density maxima, all the grid points are sorted in density descending order. Grid points can then be fractionally weighted from the weights of its neighbors with larger density. This method depends upon the formulation of flow across that dividing surfaces between the cells of two neighboring grid points, and can be applied to uniform or non-uniform grids.

We perform tests on model charge density constructed from Gaussian functions. For a one-dimensional charge density, the weight method shows that the atomic integration error

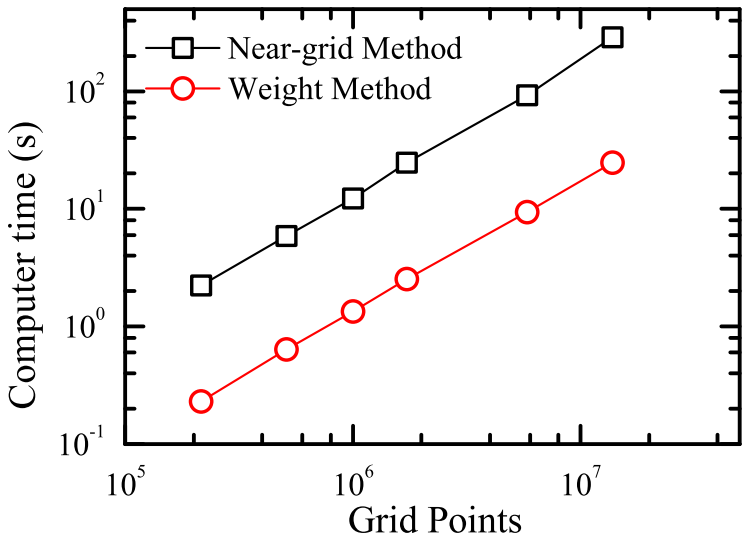

Figure 4.12: Computer time required to analyze the charge density grid for an eight-atom NaCl cell. The calculations were performed using an Intel Core2 Quad CPU Q6600, with a clockspeed of 2.40 GHz. The computer time scales linearly with respect to the number of charge density grid sizes with the weight method, as with the near-grid method. The weight method has a smaller prefactor than the near-grid method.

is inversely proportional to the square of grid points, while the integration error is inversely proportional to grid points using the near-grid method. For a three-dimensional charge density in FCC cell, the weight method shows that the atomic integration error is inversely proportional to the 2/3 of grid points, while the integration error is inversely proportional to the 1/3 power of grid points using the near-grid method. We also perform tests on more realistic systems, such as TiO_2 bulk and NaCl crystal. In both cases, the weight method reports maximal atomic integration error at least one order of magnitude smaller than the near-grid method systematically. Furthermore, we calculate the Bader charge of Na in NaCl crystal using these two methods, both give monotonic, and smooth convergence with respect to the increasing grid sizes, while they converge to slight different values, by 0.003 e. The weight method is more accurate than the near-grid method that require very fine grids.

The weight method is robust, efficient with $O(N)$ computing time, similar to the near-grid method. The former has smaller prefactor than the later, and is thus more efficient. Density

grids are evaluated from the density functional theory calculations of large molecules or materials. The weight method can be applied accurately to large systems including hundreds of atoms.

Chapter 5

Spatial extent of surfaces and defects determined from the energy density

The energy density method can provide the formation energies for more than one point defect, surface or interface in a single calculation, as well as a picture of the distribution of the energy among the surrounding atoms. By the time the energy density method was proposed, Chetty and Martin showed that the surface energy for a crystal can be calculated by an integral over a region to high-symmetry planes within the bulk of the crystal, where symmetry ensures that the gauge dependent terms integrate to zero. Therefore, two polar surface energies such as the (111) and the $(\bar{1}\bar{1}\bar{1})$ surfaces of a zincblende semiconductor GaAs can be integrated independently in a single calculation [15, 50]. Rapcewicz *et al.* [39] generalized the method to low-symmetry system such as the (0001) surface of GaN and the (0001) surface of SiC by introducing Voronoi polyhedra for the integration volumes; however, it should be noted that Voronoi polyhedra are gauge-independent integration volumes only in specific situations. Ramprasad [51] extended the application of energy density method from surfaces to point defects of metals and presented two applications on the monovacancy of Al and the (001) surface of Al. All these calculations are based on norm-conserving pseudopotential (NCPPs) with local density approximation (LDA).

To verify our implementation of the energy density method for the projector-augmented wave method (PAW) and highlight the new information it reveals, we perform DFT calculations with VASP on the atomic displacement in GaAs, GaAs(110) surface, Si monovacancy and O interstitial in Ti. We integrate the energy densities around the defect regions, and compare the integrated defect energies with values given by total energy calculations and experiments. Finally, the convergence of the atomic energy to bulk values within a *single*

calculation shows the convergence (or lack of) for each calculation.

5.1 Two atoms displacements in GaAs

Atom displacement in GaAs is a simple test on atomic energy variation from its bulk value due to geometrical distortion. The GaAs bulk is in the Zincblende structure. As shown in Figure 5.1, atoms Ga_1 and As_1 have small displacements along [111] direction by $(-0.01a, -0.01a, -0.01a)$, $(0.01a, 0.01a, 0.01a)$ from ideal positions, where a is lattice constant. Our DFT calculations are performed with the PAW method [25], with the local density approximation (LDA) [26, 20] for the exchange-correlation energy. The valence configurations for Ga is $[\mathrm{Ar}]3d^{10}4s^24p^1$ with cutoff radius 1.01 Å, and As is $([\mathrm{Ar}]3d^{10})4s^24p^3$ with cutoff radius 1.11 Å; this requires a plane-wave basis set with cutoff energy of 650 eV. This gives a lattice constant of 5.6138 Å for zincblende GaAs, compared with the experimental lattice constant of 5.65 Å. The supercell contains 4 pairs of GaAs atoms with one pair of Ga and As atoms displaced away from each other. We use Monkhorst-Pack k-point meshes [52] of $8 \times 8 \times 8$ for eight-atom cells; Brillouin-zone integration uses Gaussian smearing with $k_\mathrm{B}T = 0.1\,\mathrm{eV}$ for electronic occupancies, and the total energy extrapolated to $k_\mathrm{B}T = 0\,\mathrm{eV}$. We represent the charge density and compute energy densities on a set of grids ranging from 60^3 to 120^3 grid points.

In bulk GaAs, Ga and As atoms are 0.605 e positive and negative charged, respectively, calculated from Bader volume integration. The Bader volume of As atom is shown in middle plot in Figure 5.1. After small displacements, Bader charges on Ga_1, As_1 atom change to $+0.593\,\mathrm{e}$, $-0.615\,\mathrm{e}$ (see Figure 5.2). Figure 5.3 (top) shows the atomic energy variation from bulk energy with respect to various grid sizes. The atomic integration error is defined as the Bader volume integration of the Laplacian of electron density, which should vanish from definition. We also calculate the charge neutral volume integration of Coulomb energy density gauge term, which is usually one order of magnitude smaller than the Bader volume

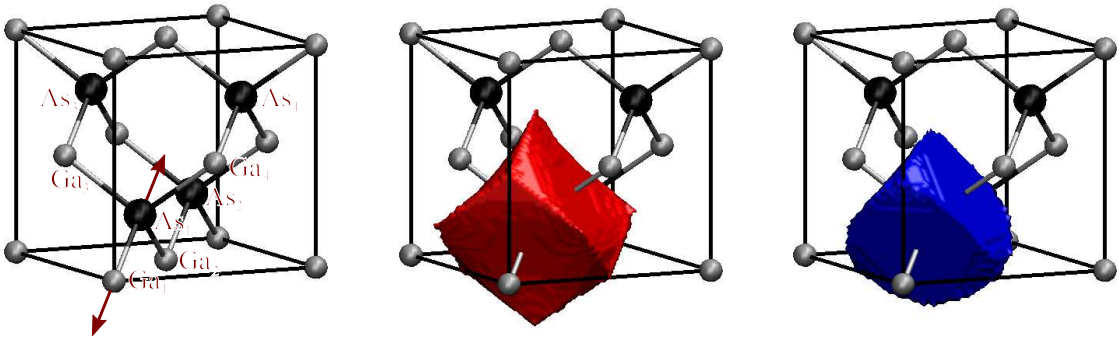

Figure 5.1: Bader volume and charge neutral volume of As in GaAs bulk. Left plot shows an eight-atom GaAs Zincblende cell. Ga_1 and As_1 atoms are displaced by $(-0.01a, -0.01a, -0.01a)$, $(0.01a, 0.01a, 0.01a)$ from ideal positions. a is lattice constant. The Bader volume (middle) has larger volume than the charge neutral volume (right) of As atom.

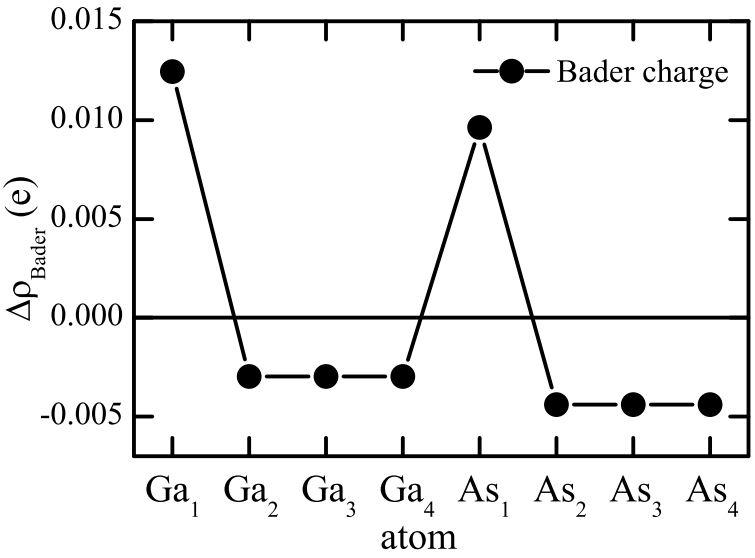

Figure 5.2: Bader charge variation on GaAs due to atomic displacements. Bader charge for each atom varies from its bulk value where Ga atom is 0.605 e positive charged and As atom is 0.605 e negative charged. Electrons move to Ga_1, As_1 atoms from other 6 atoms. Bader charges on Ga_1, As_1 atom change to $+0.593$ e, -0.615 e.

integration of the kinetic energy density gauge term. Error bars denote the atomic integration error including above two errors. Atomic energy converge on finer grid sizes with smaller atomic integration errors. Calculation on a grid of 120^3 report the atomic energy increasing

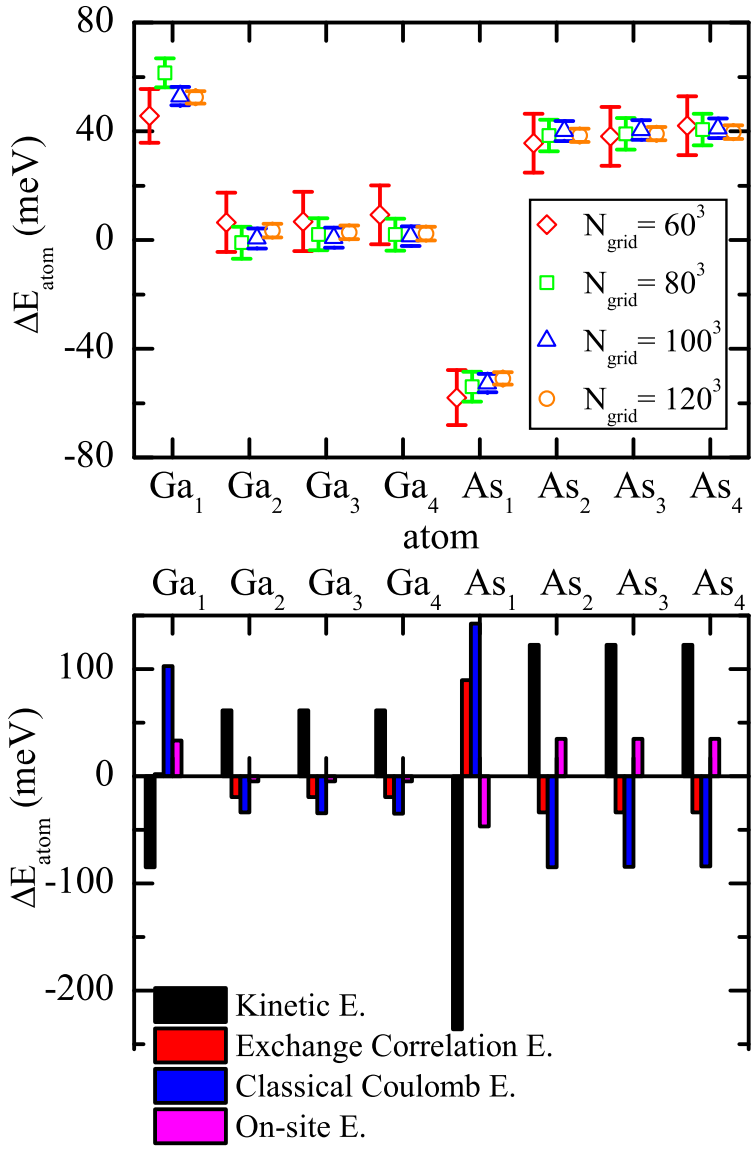

Figure 5.3: Atomic energy variation on GaAs due to atomic displacements. Atomic energy variations from bulk energy are given with respect to a range of grid sizes. bars denote integration error. Ga_1 atomic energy increases by 57 meV, As_1 atomic energy drops by 54 meV of 120^3 grid points calculation. Atomic energy breakdown into the contributions from the kinetic energy, the exchange correlation energy, the classical coulomb energy and the on-site energy on grid size 120^3 is given in the bottom plot. The kinetic energies of atoms Ga_1 and As_1 drop, and of other atoms increase. The classical Coulomb energies of atoms Ga_1 and As_1 increase, and of other atoms drop.

on Ga_1 by 57 meV, and dropping on As_1 by 54 meV. Figure 5.3 (bottom) shows the atomic energy breakdown into individual terms on the grid of 120^3. The kinetic energies of atoms Ga_1 and As_1 drop, and of other atoms increase. This agrees with the kinetic energy evaluated from Thomas-Fermi model of non-interacting electron gas, $T_{\rm TF} \propto \int \rho(\mathbf{r})^{5/3}\, d\mathbf{r}$ using the averaged Bader charge. The classical Coulomb energies of atoms Ga_1 and As_1 increase, and of other atoms drop.

5.2 GaAs(110) surface

GaAs(110) surface contains equal numbers of Ga and As atoms, and is thus a stoichiometric or non-polar surface. In traditional total energy calculation, the surface energy $\gamma_{\rm surf}$ of a stoichiometric slab is

$$\gamma_{\rm surf} = \frac{1}{2A}\left(E_{\rm slab} - N_{\rm slab}\frac{E_{\rm bulk}}{N_{\rm bulk}}\right), \tag{5.1}$$

for surface area A, where $E_{\rm slab}$ is the total energy of a GaAs slab with $N_{\rm slab}$ pairs of GaAs atoms, and $E_{\rm bulk}$ is the total energy of GaAs bulk with $N_{\rm bulk}$ pairs of atoms. Our DFT calculations are performed with the PAW method, [25] with the local density approximation (LDA) [26, 20] for the exchange-correlation energy. The valence configurations for Ga is $[\mathrm{Ar}]3d^{10}4s^24p^1$ with cutoff radius 1.01 Å, and As is $([\mathrm{Ar}]3d^{10})4s^24p^3$ with cutoff radius 1.11 Å; this requires a plane-wave basis set with cutoff energy of 650 eV. This gives a lattice constant of 5.6138 Å for zincblende GaAs, compared with the experimental lattice constant of 5.65 Å. The supercell contains 11 layers of atoms with a pair of GaAs atoms on each layer, and a vacuum gap of 8 Å to prevent the interaction between slabs under periodic boundary conditions. We use Monkhorst-Pack k-point meshes [52] of $8\times8\times8$ for bulk eight-atom cells, and $8\times8\times1$ for the slab supercell; Brillouin-zone integration uses Gaussian smearing with $k_{\rm B}T = 0.1$ eV for electronic occupancies, and the total energy extrapolated to $k_{\rm B}T = 0$ eV. We represent the charge density and compute energy densities on a grid of $84\times120\times560$. Geometry is optimized to reduce forces below 5 meV/Å. This gives a surface energy of

$50\,\mathrm{meV/Å^2}$; this agrees with Moll *et al.* 's value [53] of $52\,\mathrm{meV/Å^2}$, Qian *et al.* 's value [54] of $57\,\mathrm{meV/Å^2}$, Choudhury *et al.*'s [55] LDA value of $50\,\mathrm{meV/Å^2}$, and the experimental value [56] of $54 \pm 9\,\mathrm{meV/Å^2}$.

Figure 5.4 shows the energy change from bulk for each layer by integrating over volumes that eliminate gauge dependence. The change in energy shows differences from bulk that are mainly confined to the first two layers; the bulk-like response of the interior layers—not just for the total energy, but also the individual contributions to the energy. Determining the size-convergence of a surface calculation with total energy alone requires computing surface energies for multiple sizes; in our case, the bulk-like behavior of our center layers indicates a small finite-size error *without* requiring multiple size calculations. We can integrate the surface energy by adding the energies from the first two layers; our surface energy is $51 \pm 1\,\mathrm{meV/Å^2}$, which agrees well with the total-energy calculation of surface energy. The error estimate is specifically for the integration error over the Bader and charge-neutral volumes. Note also that we can compute the energy of each surface independently; for surfaces with different chemistry, this allows for two surface energies to be calculated from a single supercell.

Figure 5.5 shows the Bader and charge-neutral volumes for Ga and As atoms in a $(\bar{1}10)$ plane of GaAs. The Ga and As atoms all lay in a plane, and the intersection of the surfaces show the difference between the two atom-centered volumes. The Bader volumes have zero flux of the gradients of charge density through their surfaces, and are used to integrate a unique kinetic and exchange-correlation energy. The charge-neutral volumes have zero flux with the total electrostatic field, and are used to integrate a unique classical Coulomb energy. These volumes are different also from the Voronoi volumes around each atom. The atomic volumes, like the individual components of energy, become bulk-like in the center of slab. Atoms at the free surfaces have volumes that extend into the vacuum. Besides the different surfaces, the Bader volumes of As are larger than the As charge-neutral volumes.

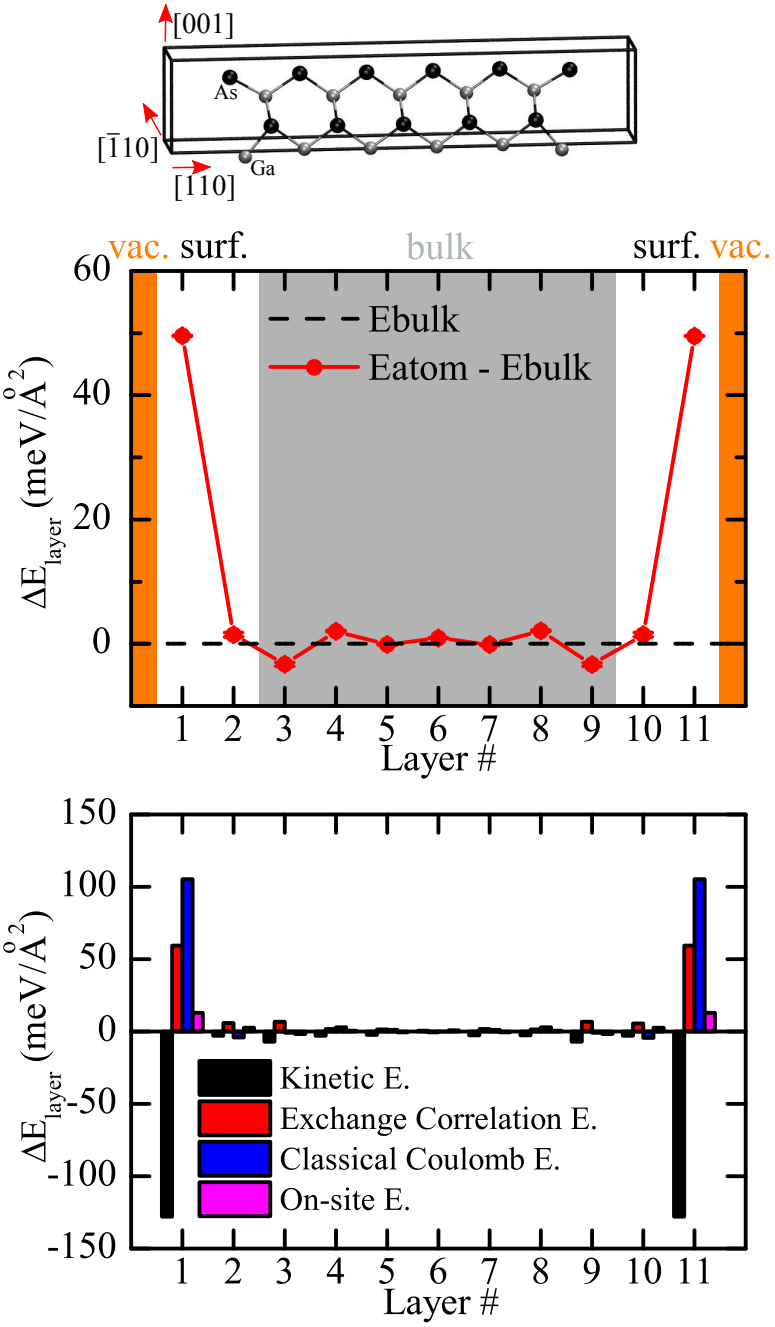

Figure 5.4: Atomic energy distribution on GaAs(110) slab. The supercell contains 11 layers of GaAs. The energy density integrated over each atomic layer divided by surface area gives the energy per layer referenced to bulk value ΔE_{layer}. The atomic integration errors are smaller than 1 meV/Å^2. Surface energy is confined to first two layers. The bulk-like behavior of the center layers indicates the sufficient thickness of slab calculation. The individual energy term contributions to each layer is shown in bottom plot. All the energy terms are bulk-like for the center layers of the slab, not just the sum. At the surface, kinetic energy decreases as the valence charge density spreads out into the vacuum.

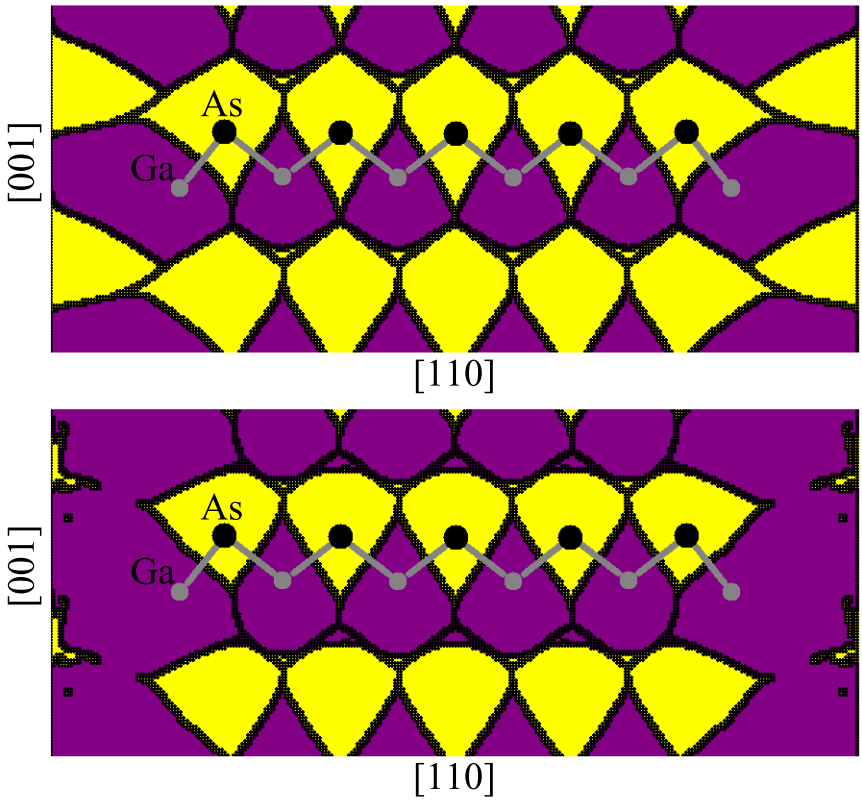

Figure 5.5: (Top) Bader volumes and (bottom) charge-neutral volumes for Ga (purple) and As (yellow) in a $(\bar{1}10)$ plane of GaAs. These integration volumes define the unique atomic kinetic and exchange-correlation energies, and classical Coulomb energies. Each volume contains a single atom, but the two volumes are different for the same atom.

5.3 Si point defect

Silicon is a most important semiconductor. The vacancy and self-interstitial in bulk Si have been studied in theoretical and experimental reseach for several decades. From total energies, the formation energy of a vacancy ΔH_v is

$$\Delta H_v = E_v^{N-1} - \frac{N-1}{N} E^N, \tag{5.2}$$

and the formation energy of a self-interstitial ΔH_i is

$$\Delta H_i = E_i^{N+1} - \frac{N+1}{N} E^N, \tag{5.3}$$

where E_v^{N-1}, E^N, and E_i^{N+1} are the total energy of the $N-1$ atom supercell with one vacancy, the N atom perfect lattice supercell, and $N+1$ atom supercell with one interstitial, respectively.

Our DFT calculations are performed with the PAW method with the generalized-gradient approximation (GGA) of Perdew and Wang (PW91) [21] for the exchange-correlation energy. The valence configurations for Si is [Ne]$3s^2 3p^2$ with cutoff radius 1.01 Å; this requires a plane-wave basis set with cutoff energy of 417 eV. This gives a lattice constant of 5.4674 Å for diamond Si, compared with the experimental lattice constant of 5.43 Å. Monovacancy calculation uses a $2 \times 2 \times 2$ simple cubic supercell with a vacancy containing 63 atoms, and self-interstitial calculation uses a $3 \times 3 \times 3$ simple cubic supercell with an interstitial containing 217 atoms. We use a $8 \times 8 \times 8$ Monkhorst-Pack k-point mesh for 8-atom cells, a $4 \times 4 \times 4$ Monkhorst-Pack k-point mesh for 64-atom cells and 216-atom cells; Brillouin-zone integration uses Gaussian smearing with $k_{\mathrm{B}}T = 0.15\,\mathrm{eV}$, and the total energy extrapolated to $k_{\mathrm{B}}T = 0\,\mathrm{eV}$. Geometry is optimized to reduce forces below 5 meV/Å.

5.3.1 Monovacancy

We obtain the formation energy of 3.65 eV from the total energy calculation. Our results can be compared with several other recent DFT calculations. Wright [57] performed LDA [20] and GGA-PBE [22] calculations in 215, 511-, and 999-atom supercells to get formation energies of 3.53 eV, 3.49 eV, and 3.47 eV with LDA and 3.66 eV, 3.63 eV, and 3.62 eV with GGA. Puska *et al.* [58] performed LDA calculations in 31-, 63-, 127-, and 215-atom supercells to get formation energies of 3.98 eV, 3.42 eV, 3.44 eV, and 3.31 eV. Experiments have found a formation energy of 3.6 ± 0.2 eV. [59].

We perform energy density calculation on a grid of $200 \times 200 \times 200$. Figure 5.7 shows the energy change from bulk for shells surrounding a Si vacancy. The primary contribution to the vacancy formation energy comes from the first five shells, becoming bulk-like at larger distances. Summing the atomic energies up to the fifth shell gives a formation energy of 3.57 ± 0.05 eV, which is similar to the total energy calculation. The importance of the fifth shell over the third and fourth shells can also be seen in charge disturbances from a vacancy. Kane [60] showed charge disturbances around a Si monovacancy out to the 27^{th} shell, while the first two shells contribute 60% of the charge disturbance. The most striking future was charge concentrates on the $\{110\}$ planar zigzag chains of atoms, such as [000], [111], [220], [331], [440], as so on, in the units of $a/4$. Twelve such chains exist by symmetry. After reaching the fifth shell at $\frac{a}{4}\langle 331 \rangle$, the charge decays monotonically along the coplanar chains. Kane connected this result to the importance of the fifth-neighbor interaction in the valence force model [61] of covalent phonon spectra. The valence force model is an empirical model connecting force constants to the electronic configuration. The fifth-neighbor interaction is proportional to $r^2 \Delta\varphi \Delta\varphi'$ that affects bond angles φ, φ' with a common bond along a zigzag chain. For example, two angles φ, φ' are defined by the sequence of atomic sites [000], [111], [220], and [111], [220], [331], and r is the length of common bond between atomic site [111] and [220]. The fifth neighbor has a stronger interaction with the bond-bending than third

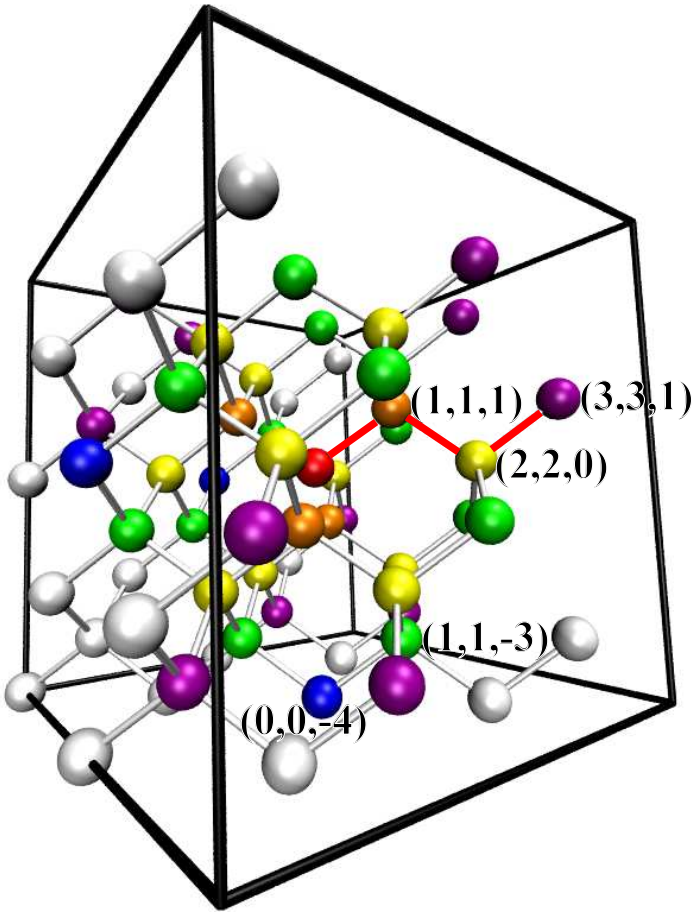

Figure 5.6: Si monovacancy (red) in $2 \times 2 \times 2$ simple cubic supercell. The first five shells are $\frac{a}{4}\langle 111\rangle$ orange, $\frac{a}{4}\langle 220\rangle$ yellow, $\frac{a}{4}\langle 11\bar{3}\rangle$ green, $\frac{a}{4}\langle 00\bar{4}\rangle$ blue, and $\frac{a}{4}\langle 331\rangle$ violet. Only three shells have displacements greater than 0.01 Å: first shell with 4 atoms relax inward by 0.17 Å; the second shell with 12 atoms, relax inward by 0.05 Å; and the fifth shell relax inward by 0.02 Å. These three shells form a zigzag chain (red) from the vacancy with a strong interaction.

and fourth neighbors; we see a similar change in energy for the vacancy.

5.3.2 Tetrahedral self-interstitial

Many different Si self-interstitial configurations have been proposed. The most widely studied are the hexagonal, tetrahedral, and split-$\langle 110\rangle$ interstitials as they are low in energy. Figure 5.8 shows a tetrahedral interstitial which has larger interatomic distance compared with the other two interstitials. This would require small energy cutoff for calculation. In

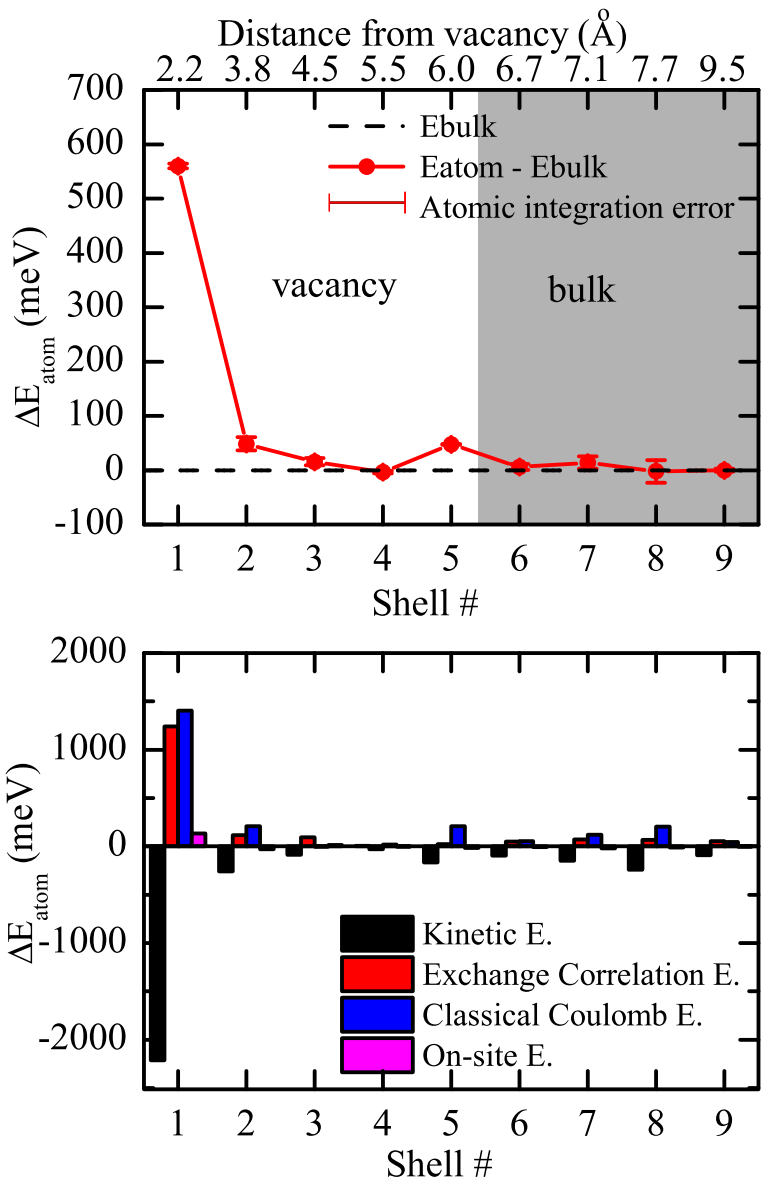

Figure 5.7: Atomic energies of Si in $2 \times 2 \times 2$ simple cubic supercell with monovacancy. The atomic energy confirms a strong interaction along a zigzag chain (red) with energies of 560 ± 5 meV (first shell), 49 ± 12 meV (second shell), 47 ± 1 meV (fifth shell); compared with 16 ± 7 meV (third shell) and -4 ± 3 meV (fourth shell). As with a free surface, the kinetic energy drops close to the vacancy due to decreasing of valence charge density.

this work, we study tetrahedral self-interstitial, and obtain the formation energy of 4.02 eV from the total energy calculation. Our results can be compared with several other recent DFT calculations. Centoni *et al.* performed PW91-GGA [21] calculation using the ultra-soft Vanderbilt-type [24] pseudopotential in 257-atom supercell to get formation energy of 4.06 eV. [62] Needs performed LDA and PW91-GGA calculations using the norm-conserving pseudopotential in 129-atom supercell to get the formation energies of 3.43 and 4.07 eV. [63] Experiments can not detect the Si self-interstitial directly. The experimental study of self-diffusion obtains the diffusion activation energy, the sum of the formation energy and the migration energy for the self-interstitial of $4.68^{+0.12}_{-0.15}$ eV. [64] Assuming the migration energy of 0.1-0.3 eV [63], the self-interstitial formation would be above 4.23 eV.

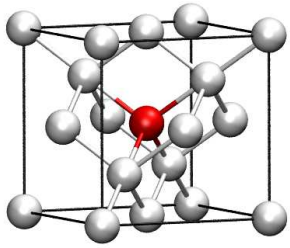

Figure 5.8: Si tetrahedral self-interstitial in conventional cubic cell. The high-symmetry tetrahedral site (red) has four nearest-neighbor atoms as well as six next nearest-neighbor atoms.

We perform energy density calculation on a grid of $180 \times 180 \times 180$. Figure 5.9 shows the energy change from bulk for shells surrounding a Si self-interestitial. The primary contribution to the interestitial formation energy comes from the first two shells, however atomic energy decays slowly to the bulk value due to long-ranged elastic interaction. Figure 5.10 presents the integrated formation energy from interstitial site to individual shells. As atomic energy decays monotonically to bulk value after 11^{st} shell, we fit the formation energy from 12^{th} shell to the outest shell using equation $H_i = 4.07 - 7.7 \times 10^6 \, r^{-7.8}$, where r is the distance

from the interstitial site. We obtain the self-interstitial formation energy of 4.07 eV.

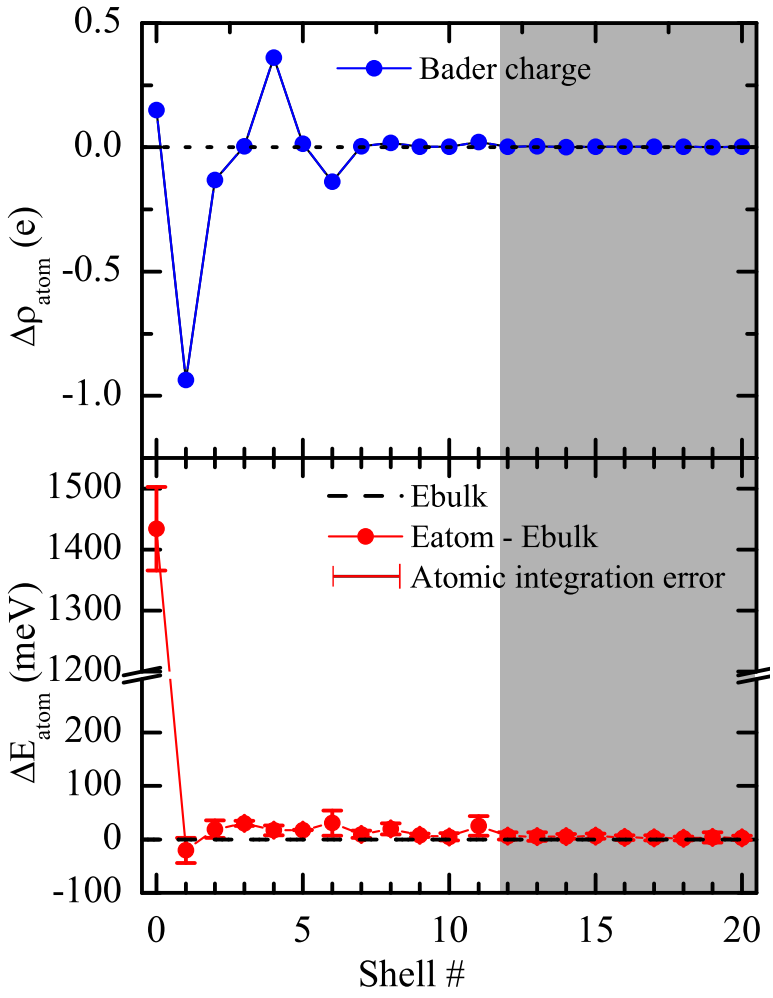

Figure 5.9: Bader charges and Atomic energies surrounding a Si self-interstitial. Bader charges and atomic energies oscillate, and monotonically decay to bulk value after 11st shells.

5.4 HCP Ti with O interstitial

In this section, we consider the formation energy of an oxygen interstitial in the octahedral site of hexagonal close-packed (HCP) titanium. The octahedral site is more stable than

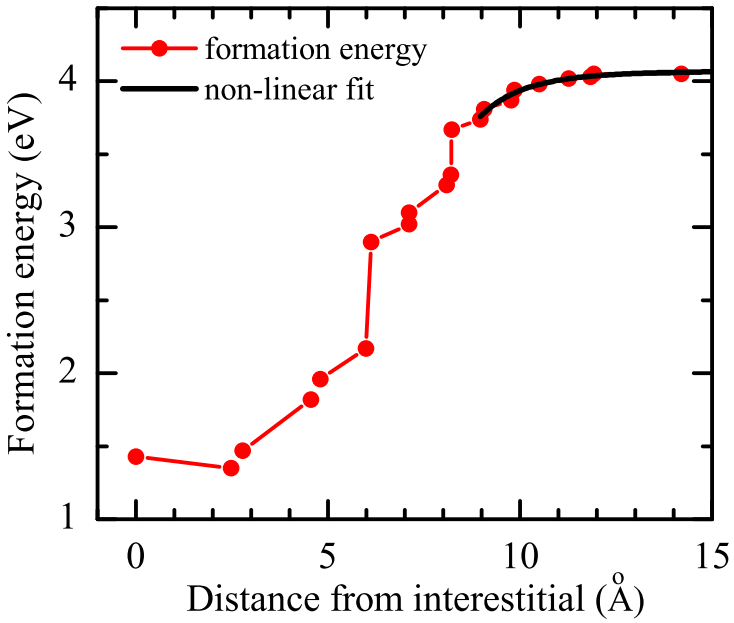

Figure 5.10: Si self-interstitial formation energy. Formation energy H_i sums up from the interstitial site shell by shell. Non-linear fit of formation energy from 12nd shell with respective to distance from interstitial using equation $H_i = 4.07 - 7.7 \times 10^6\, r^{-7.8}$ gives the formation energy of 4.07 eV.

tetrahedral or hexahedral interstitial site [65]. The formation energy is

$$E_{\mathrm{i}}^{\text{Ti-O}} = E(\mathrm{Ti} + \mathrm{O_i}) - E(\mathrm{Ti}) - \frac{1}{2}E(\mathrm{O_2}), \tag{5.4}$$

where $E(\mathrm{Ti} + \mathrm{O_i})$ and $E(\mathrm{Ti})$ are the total energy of relaxed supercells with and without an oxygen atom, and $E(\mathrm{O_2})$ is the total energy of oxygen molecule. Our DFT calculations are performed by use of the PAW method with the GGA-PW91 [21] for the exchange-correlation energy. The valence configurations for Ti is $[\mathrm{Ne}]3s^2 3p^6 4s^2 3d^2$ with cutoff radius 1.22 Å, and O is $[\mathrm{He}]2s^2 2p^4$ with cutoff radius 0.80 Å; this requires a plane-wave basis set with cutoff energy of 500 eV. This gives a lattice constant of a =2.933 Å, c =4.638 Å, and c/a =1.581 for HCP Ti, compared with the experimental lattice constant of a=2.951 Å, c =4.684 Å, and c/a =1.587. [66] The supercell contains 96 Ti atoms ($4 \times 4 \times 3$) and 1 O atom. We use a

2×2×2 Monkhorst-Pack k-point mesh; Brillouin-zone integration uses the Methfessel-Paxton method with $k_{\mathrm{B}}T = 0.1\,\mathrm{eV}$ for electronic occupancies, and the total energy extrapolated to $k_{\mathrm{B}}T = 0\,\mathrm{eV}$. We represent the charge density and compute energy densities on a grid of $180 \times 180 \times 216$. Geometry is optimized to reduce forces below $20\,\mathrm{meV/Å}$. This gives an oxygen interstitial formation energy of –6.19 eV, with a nearest-neighbor distance between Ti and O of 2.08 Å. Hennig *et al.* performed GGA-PW91 calculations using ultrasoft Vanderbilt-type [24] pseudopotentials in the same supercell to get a formation energy of –6.12 eV and nearest-neighbor distances of 2.06–2.09 Å. [65]

Figure 5.11 shows the calculated Ti atomic energy change from bulk value for each shell. The change in energy shows the differences from bulk are mainly confined to the first two shells, with bulk-like behavior for shells further away from the oxygen atom. The energy density and charge density oscillates and decays away from the interstitial. They peak at the 6th shell and the 13th shell with a wavelength of 1.9 Å. Weiss's Compton profile [67] measured the Fermi momentum of Ti as 1.08 ± 0.06 a.u., which corresponds to a Friedel oscillation wavelength of 1.5 Å. Jepson's [68] earlier calculation using linear muffin-tin-orbital method obtained the Fermi energy of Ti as 0.667 Ryd which corresponds to the Friedel oscillation wavelength of 2.0 Å. Adding the atomic energy change of first two shells from the O interstitial and the atomic energy change of O atom, we obtain the interstitial formation energy of -6.13 ± 0.01 eV, which agrees with the total-energy calculation.

5.5 Summary

In this chapter, we apply the energy density method and gauge independent integration algorithm to studying surfaces and point defects. Different systems have been studied from semiconductors to metals. One calculates surface energies by integrating the energy density in the surface region; and defect energies of mono-vacancy or mono-interstitial by integrating the energy density in the point defect region. Our calculation results of defect energies by

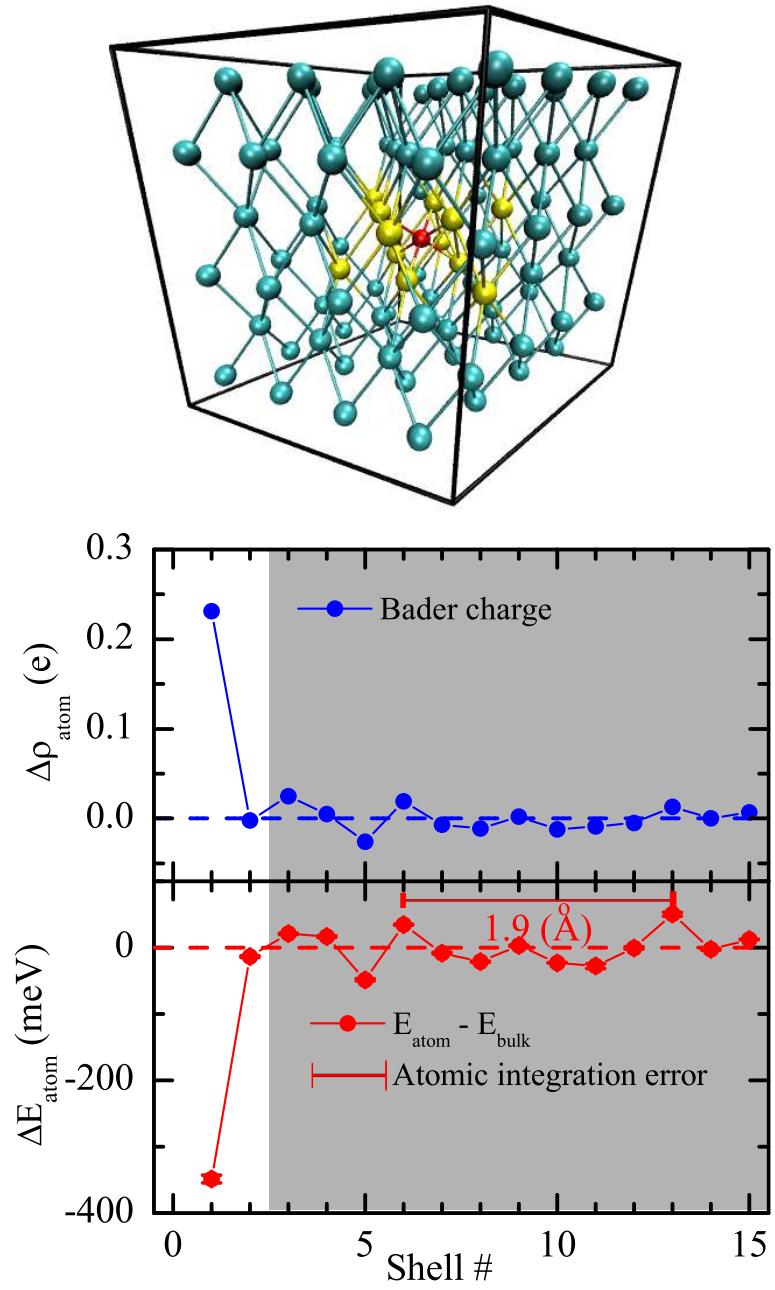

Figure 5.11: HCP Ti $4 \times 4 \times 3$ supercell with an O interstitial (red) in an octahedral site. The formation energy is confined to the first two Ti shells (yellow); away from the oxygen atom, the charge density and energy density of Ti atoms experience Friedel-like oscillations. There is a charge transfer of 1.45 e to the interstitial O atom.

using the energy density method are comparable with experimental values and total energy calculations. The energy density also produces some information that cannot be found from

total energy calculations.

In the calculation of GaAs(110) surface, the atomic energy distribution on GaAs(110) slab shows the energy difference from the bulk is confined to the first two layers, which is the type of information that cannot be found from total energy calculations alone; the extent of the surface region shows the convergence of the calculation with slab size. This replaces the conventional convergence tests on different size slabs to find out what thickness is sufficient. Without thickness convergence test, The bulk-like behavior of our slab center layers indicates the sufficient thickness of current slab calculation. In the calculation of Si monovacancy, the atomic energy variation from bulk value concentrates mainly on the first five shells. The similar phenomena of charge density has been reported, and is relavent to the fifth-neighbor interaction between atoms on the {110} planar zigzag chains. The O interstitial in the hexagonal close-packed (HCP) Ti crystal demonstrates a Friedel oscillation both in charge density and energy density.

Chapter 6

Stability and properties of Au/TiO_2 interfaces

As reviewed in Section 1.1, two different epitaxies of Au nanoparticles have been observed on $TiO_2(110)$ surfaces: Au(111) and Au(100) in HRTEM experiment, Au(111) epitaxy is more frequently observed than Au(100). Using HAADF-STEM, a specially reconstructed interface has been examined for epitaxial Au(111) sitting on reconstructed $TiO_2(110)$ 1×2 surface. [11] On the theoretical side, density functional theory calculations have studied the optimum size and stable adsorption of Au nanoparticles on rutile $TiO_2(110)$. Single Au atom is energetically favorable on the atop site above five-fold coordinated (5c) Ti atom on a stoichiometric TiO_2 surface, [69] and is most stable on the two-fold coordinated (2c) bridging O vacancy site on a reduced surface. [70, 71, 72] O vacancies cause a stronger binding of Au atoms [73], nanoclusters [74, 75, 76], or nanorows [75] to the reduced TiO_2 surface than to the stoichiometric surface. Apart from the stoichiometric and reduced TiO_2 surfaces, Shi *et al.* found the O-rich interface is the most stable at low temperature of catalytic reaction after examining the Au-rod/$TiO_2(110)$ in the orientation Au(111)//$TiO_2(110)$ with different interface stoichiometry and various rigid-body translations. [77] Recently, Shibata *et al.* used the PAW method [32] with PBE-GGA exchange-correlation energy [22] to examine two and nine Au(110) atomic layers supported on reduced $TiO_2(110)$, and demonstrated that both the atomic and electronic structure of two-layer Au are reconstructed, while the lattice coherency decays rapidly across the interface for nine-layer Au. [78] We are the first to compare different Au/TiO_2 interfaces: Au(111)//$TiO_2(110)$ and Au(100)//$TiO_2(110)$, with and without bridging oxygen, Au(111) on added-row 1×2 $TiO_2(110)$ reconstruction [79], and Au(111) on a new proposed TiO 1×2 reconstruction. The density functional theory

energy density method evaluates the atomic energy, observes the change in atomic energy during forming interface, and computes the work of adhesion to determine the equilibrium interfacial structure. The comparison with experimental geometry allows us to validate our predicted structures. Before analyzing the interface, we study the geometric and electronic properties of the two compositions: rutile TiO_2 from bulk to (110) surfaces; FCC Au from bulk to (111) and (100) surfaces.

6.1 TiO_2 surface structure

6.1.1 Bulk structure

Two mainly applied and widely studied structures of titanium dioxide crystalline are rutile and anatase. They both are in tetragonal unit cell. Figure 6.1 shows a tetragonal unit cell of rutile TiO_2 containing two Ti atoms and four O atoms. Each Ti atom is surrounded by six O atoms in a distorted octahedral configuration.

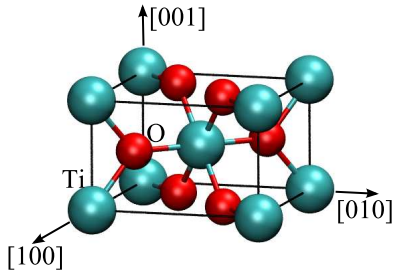

Figure 6.1: Rutile TiO_2 structure. The experimental [6] lattice constants are $a = 4.584\,\text{Å}$; $c = 2.953\,\text{Å}$; $u = 0.305\,\text{Å}$. Cyan dots are Ti atoms, red dots are O atoms.

We perform DFT calculation on TiO_2 bulk by use of the PAW method and USPPs [24] with the LDA, PW91-GGA, and PBE-GGA for the exchange-correlation energy. The calculated structural properties of bulk TiO_2 are listed in Table 6.1 as well as experimental data. We use Monkhorst-Pack k-point meshes of $4 \times 4 \times 6$ for six-atom cell; Brillouin-zone integration uses Gaussian smearing with width $k_B T = 0.1\,\text{eV}$ for electronic occupancies, and the

total energy extrapolated to $k_BT = 0\,\text{eV}$. Both LDA and GGA ultrasoft pseudopotentials introduce too large atomic cutoff radius. The calculated nearest neighboring distance 1.94 Å is smaller than the sum of two atomic radii, this would cause atomic core overlap. Since we do not consider the core overlap term in energy density, we will only use the PAW method for our later calculations.

	R_{Ti}(Å)	R_{O}(Å)	a(Å)	c(Å)	u	$d_{n.n.}^{\text{Ti-O}}$(Å)
expt [6]:			4.584	2.953	0.305	
PAW-GGA-PW91:	1.217	0.582	4.645	2.969	0.305	1.96
PAW-GGA-PBE:	1.217	0.582	4.649	2.970	0.305	1.96
PAW-LDA:	1.217	0.582	4.561	2.926	0.304	1.93
USPPs-LDA:	1.476	1.005	4.552	2.925	0.304	1.94

Table 6.1: Structural properties of rutile TiO_2. Calculations are performed using different DFT method or exchange-correlation functionals. In all of the PAW calculations, the valence configurations for Ti is [Ne]$3s^2 3p^6 4s^2 3d^2$ with the semi core s states treated as valence, and O is [He]$2s^2 2p^4$ with hard pseudopotential, so that nearest neighboring distance $d_{n.n.}^{\text{Ti-O}} > R_{\text{Ti}} + R_{\text{O}}$, no overlap exists between two atomic cores. Ultrasoft pseudopotentials have too large atomic cutoff radii.

6.1.2 Rutile $TiO_2(110)$ 1×1 surface

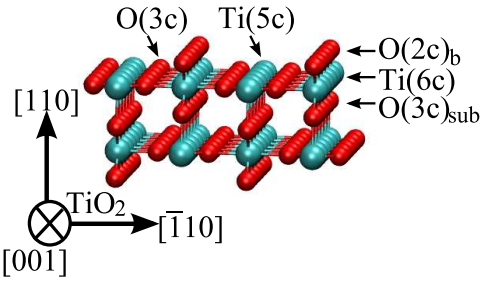

Figure 6.2: Geometry of rutile $TiO_2(110)$ surface. Atomic coordination number is given in parenthesis. Rutile (110) surface contains two different kinds of Ti atoms, five-fold coordinated (5c) Ti atoms with one dangling bond perpendicular to (110) surface and six-fold coordinated (6c) Ti atoms. Each Ti (6c) atoms is surrounded by four in-plane O (3c) atoms, one bridging O $(2c)_b$ atom, and one subsurface bridging O $(3c)_{sub}$ atom.

The rutile (110) surface shown in Figure 6.2 is the most stable TiO_2 1×1 surface. The bridging O atoms are undercoordinated, thus would be removed relatively easily by thermal annealing, and lead to surface defects. Kiejna *et al.* performed GGA calculations using USPPs and PAW to get surface energies of $36\,\mathrm{meV/Å^2}$ and $29\,\mathrm{meV/Å^2}$. [80] Ramamoorthy *et al.* performed LDA calculations using USPPs in 3- to 6-layer supercells to get surface energies falling in the range of $69\,\mathrm{meV/Å^2}$ to $52\,\mathrm{meV/Å^2}$. [81]

		$\gamma(\mathrm{meV/Å^2})$		
Slab L	Relaxed L	PW91-GGA	PBE-GGA	LDA
5	1-5	33.4	31.7	57.4
6	unrelaxed	85.1	84.0	107.7
	1,2	32.7	31.0	56.5
	1-6	23.8	21.7	49.6
7	unrelaxed	84.7	83.9	107.7
	1,2	33.1	31.1	56.6
	1,2,3	29.9	27.9	54.0
	1-7	29.6	27.7	54.1
8	unrelaxed	85.1	83.8	107.7
	1,2	32.7	31.0	56.6
	1,2,3	29.4	27.6	53.9
	1,2,3,4	27.6	25.5	52.3
	1-8	24.4	22.0	50.6

Table 6.2: Surface energy of partially or completely relaxed TiO_2(110). Calculations are performed by using PAW-PW91-GGA; PAW-PBE-GGA; PAW-LDA.

Table 6.2 shows the energy convergence of the unrelaxed, partially and completely relaxed surfaces with respect to the thickness of the slabs. Supercells range from 5- through 8-trilayers of atoms with two Ti and four O atoms on each trilayer (see right figure in Table 6.3 as a example of 7-trilayer slab), and a vacuum gap of 10.5 Å to prevent the interaction between slabs under periodic boundary conditions. Energies of unrelaxed TiO_2(110) surface converge to $85\,\mathrm{meV/Å^2}$ (GGA-PAW91), $84\,\mathrm{meV/Å^2}$ (GGA-PBE), and $108\,\mathrm{meV/Å^2}$ (LDA). The geometry relaxations greatly reduce surface energies to $33-24\,\mathrm{meV/Å^2}$ (GGA-PAW91), $32-22\,\mathrm{meV/Å^2}$ (GGA-PBE), and $57-50\,\mathrm{meV/Å^2}$ (LDA). The odd and even layer slab exhibit nonequivalence effect on surface energy, this is caused by the displacements of Ti

and O atoms on the surface layers, and the presence/absence of mirror symmetry of odd/even layer slab. Partially relaxed slab presents convergence by increasing the number of relaxed layer by one each time. For example, The surface energy of relaxed side of 8-trilayer slab gradually reduces from 31 to 25 meV/Å^2 in the PAW-PBE calculation that the number of relaxed layer increases from 2 to 4 by one each time. Interestingly, the nonequivalence effect between odd and even layer slabs doesn't exist for partially relaxed slabs. The relaxed surface energy calculated from 6-, 7-, 8-trilayer slab with only left 2-trilayer(or 3-trilayer) relaxed report very close values.

	Experiment		Theory[GGA-PW91]	
Atom	SXRD [82]	LEED-IV [83]	USPP [80]	PAW
1:Ti(6c)	0.12 ± 0.05	0.25 ± 0.03	0.21	0.28
2:Ti(5c)	-0.16 ± 0.05	-0.19 ± 0.03	-0.21	-0.14
3:O$(2c)_b$	-0.27 ± 0.08	0.10 ± 0.05	0.00	0.07
4,5:O	0.05 ± 0.05	0.27 ± 0.08	0.14	0.21
4,5:O[$\bar{1}$10]	$\pm 0.16 \pm 0.08$	$\pm 0.17 \pm 0.15$	± 0.05	± 0.04
6:O	0.05 ± 0.08	0.06 ± 0.10	-0.01	0.06
7:Ti	0.07 ± 0.04	0.14 ± 0.05	0.13	0.20
8:Ti	-0.09 ± 0.04	-0.09 ± 0.07	-0.13	-0.07
9:O	0.00 ± 0.08	0.00 ± 0.08	-0.02	0.05
10,11:O	0.02 ± 0.06	0.06 ± 0.12	-0.01	0.05
10,11:O[$\bar{1}$10]	$\pm 0.07 \pm 0.06$	$\pm 0.07 \pm 0.18$	± 0.02	± 0.02
12:O	-0.09 ± 0.08	0.00 ± 0.17	-0.03	0.04
13:Ti			0.06	0.06
14:Ti			-0.06	-0.06
15:O	-0.12 ± 0.07	0.01 ± 0.13	0.02	0.01
16,17:O			0.01	0.02
16,17:O[$\bar{1}$10]			± 0.01	± 0.01
18:O			0.00	-0.01

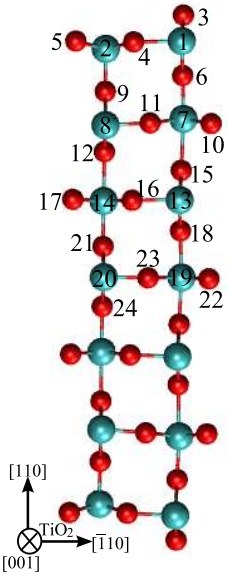

Table 6.3: Atomic relaxation on TiO_2(110) along [110] and [$\bar{1}$10] directions (in Å). Calculations are performed in this work by use of the PAW method with the generalized-gradient approximation (GGA) of Perdew and Wang (PW91) for the exchange-correlation energy for completely relaxed 7-layer slab, and compared with experimental data and others' DFT calculations. Atomic index in first row are labeled in right figure.

Geometry relaxation of atoms in TiO_2(110) slab plays an important role in the oscillation of relaxed surface energy in terms of slab thickness. We give an example of atomic

displacements in 7-trilayer completely relaxed slab as shown in the right figure in Table 6.3, and compare with data from surface X-ray diffraction (SXRD) [82] and quantitative low energy electron diffraction (LEED-IV) [83] experiments and other's USPPs-PW91 calculation in Table 6.3. Our calculation agree with LEED measurements well. Atomic displacements mainly occur on the top layer perpendicular to the (110) surface; atoms in the deeper layer also have displacements with small magnitude; in-plane O atoms exhibit displacements along $[\bar{1}10]$ direction.

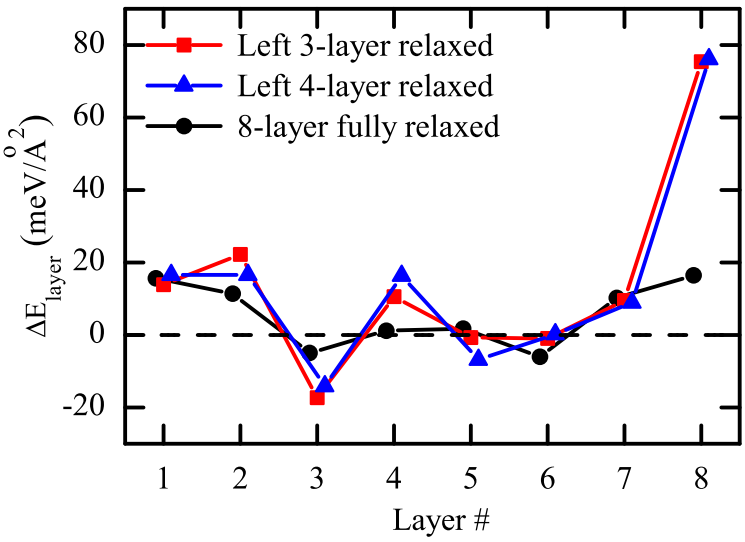

Figure 6.3: Atomic energy distribution on 8-trilayer TiO_2(110) slab. Atomic energies per trilayer with respect to the bulk value are given for partially and completely relaxed TiO_2 slabs.

We perform energy density calculation of 8-trilayer TiO_2 slab on a grid of $90 \times 40 \times 480$. Three different slab relaxations are discussed: completely relaxed, left 3-trilayer and left 4-trilayer partially relaxed slabs. Figure 6.3 shows the calculated energy change from bulk value for each trilayer of atoms. The right sides of 3- and 4-trilayer partially relaxed slabs are unrelaxed. The slab with only left 3 trilayers relaxed reaches bulk-like behavior at 5^{th} and 6^{th} trilayers. The slab with only left 4 trilayers relaxed reaches bulk-like behavior at 6^{th} trilayer. The completely relaxed slab has two center trilayers reach bulk-like behavior.

From one calculation of 3-trilayer relaxed slab, one obtains a (110) relaxed surface energy of $30 \pm 1\,\mathrm{meV/\AA^2}$ from energy density integration of left 3 trilayers and a (110) unrelaxed surface energy of $85 \pm 1\,\mathrm{meV/\AA^2}$ from energy density integration of right 2 trilayers. Similarly, from one calculation of 4-trilayer relaxed slab, one obtains a (110) relaxed surface energy of $29 \pm 1\,\mathrm{meV/\AA^2}$ from energy density integration of left 4 trilayers and a (110) unrelaxed surface energy of $85 \pm 1\,\mathrm{meV/\AA^2}$ from energy density integration of right 2 trilayers. For the completely relaxed slab, energy density integration of left (or right) 3 trilayers gives a relaxed surface energy of $22 \pm 1\,\mathrm{meV/\AA^2}$. This leads to three important conclusions: 1. the energy difference from the bulk is almost completely confined to the first 3 trilayers of 8-trilayer slab; the left 4 trilayers of 8-trilayer slab with left 3-trilayer relaxed; and the left 5 trilayers of 8-trilayer slab with left 4-trilayer relaxed. 2. Without thickness convergence test, The bulk-like behavior of our slab center trilayers indicates the sufficient thickness of current slab calculation. 3. The energy density method can provide the surface energies for more than one surface in one calculation.

6.2 Au surface structures

6.2.1 Bulk structure

We perform calculations on structural properties of FCC Au (Table 6.4) with different pseudopotentials and computational methods. The valence configurations for Au is [Xe]$6s^1 5d^{10}$; this requires a plane-wave basis set with cutoff energy of 400 eV. We use Monkhorst-Pack k-point meshes of $13 \times 13 \times 13$ for bulk four-atom cells, and $13 \times 13 \times 1$ for slab supercells; Brillouin-zone integration uses the Methfessel-Paxton method [84] with $k_\mathrm{B}T = 0.2\,\mathrm{eV}$ for electronic occupancies, and the total energy extrapolated to $k_\mathrm{B}T = 0\,\mathrm{eV}$. Geometry is optimized to reduce forces below $5\,\mathrm{meV/\AA}$. The optimized lattice constants agree with experiment data well.

	R_{Au}(Å)	a(Å)
expt:		4.08
PAW-GGA-PW91:	1.32	4.172
PAW-GGA-PBE:	1.32	4.171
PAW-LDA:	1.32	4.062
USPPs-LDA:	1.50	4.068

Table 6.4: Structural properties of FCC Au. Calculations are performed using different DFT method or exchange-correlation functionals. The valence configuration for Au is [Xe]$6s^1 5d^{10}$.

6.2.2 Au(111) surface

Table 6.5 shows the energy convergence of the completely relaxed surfaces with respect to the thickness of the slabs. Supercells range from 4- through 8-layers of atoms with two Au atom on each layer, and a vacuum gap of 12 Å to prevent the interaction between slabs under periodic boundary conditions. Our calculated surface energy converges to 43 meV/Å^2 (GGA-PAW91), 43 meV/Å^2 (GGA-PBE), and 70 meV/Å^2 (LDA); this agrees with Zólyomi *et al.* 's [85] PAW-PBE value of 44 meV/Å^2, and Swart *et al.* 's [86] USPPs-GGA value of 44 meV/Å^2 and USPPs-LDA value of 69 meV/Å^2. The LDA calculations give surface energies much larger than GGA calculations.

	γ(meV/Å^2)		
Slab L	PW91-GGA	PBE-GGA	LDA
4	44.2	44.0	71.2
5	42.4	42.4	69.5
6	43.6	43.2	70.5
7	42.7	42.6	70.6
8	43.2	42.9	70.4

Table 6.5: Surface energy of completely relaxed Au(111).

We represent the charge density and compute energy densities on a grid increasing from $48 \times 84 \times 360$ to $48 \times 84 \times 480$ for 5- to 8-layer supercells. Figure 6.4 shows the energy change from bulk value for each layer of atoms. Energy density integration of left (or right) 2 layers of 6-layer slab gives a surface energy of 40 ± 1 meV/Å^2; left (or right) 3 layers

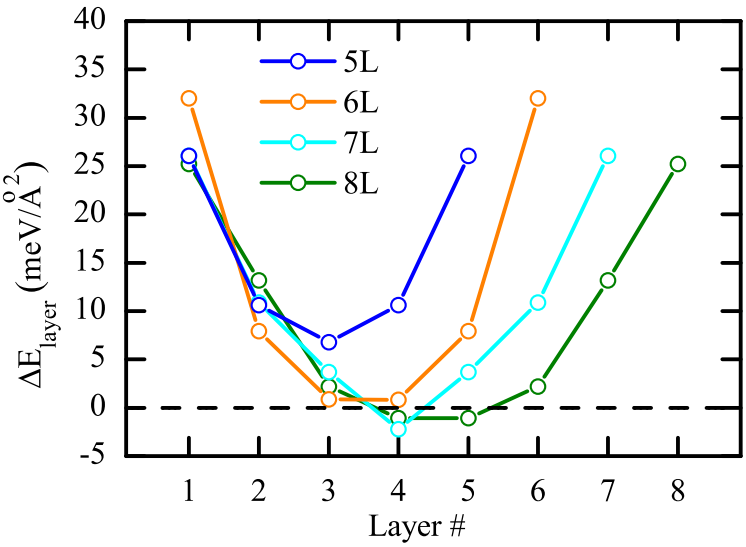

Figure 6.4: Atomic energy distribution on Au(111) slab. The atomic integration errors are smaller than 1 meV/Å^2. The bulk-like behavior of the center layer(s) of 6-, 7-, and 8-layer slabs indicates the sufficient thickness of slab calculation.

of 7-, and 8-layer slabs both give a surface energy of 41 ± 1 meV/Å^2. Although the total energy calculations present close values of surface energy on 5-, and 6- or more layer slabs, the atomic energy distribution shows the center layers of 5-layer slab have not reached the bulk-like behavior. Unlike the convergence test of the traditional total energy calculation, the required thickness of a slab can be directly determined by observing the variation of atomic energy from the bulk value.

6.2.3 Au(100) surface

Table 6.6 shows the energy convergence of the completely relaxed surfaces with respect to the thickness of the slabs. Supercells range from 4- through 7-layers of atoms with one Au atom on each layer, and a vacuum gap of 10.5 Å to prevent the interaction between slabs under periodic boundary conditions. Our calculated surface energy converges to 53 meV/Å^2 (GGA-PAW91), 52 meV/Å^2 (GGA-PBE), and 84 meV/Å^2 (LDA); this agrees with Zólyomi *et al.* 's [85] PAW-PBE value of 54 meV/Å^2, and Swart *et al.* 's [86] USPPs-GGA value

of 56 meV/Å^2 and USPPs-LDA value of 83 meV/Å^2. The LDA calculations give surface energies much larger than GGA calculations.

	γ(meV/Å^2)		
Slab L	PW91-GGA	PBE-GGA	LDA
4	53.7	53.3	84.4
5	53.1	52.7	83.9
6	52.6	52.5	83.5
7	52.6	52.3	83.5

Table 6.6: Surface energy of completely relaxed Au(100).

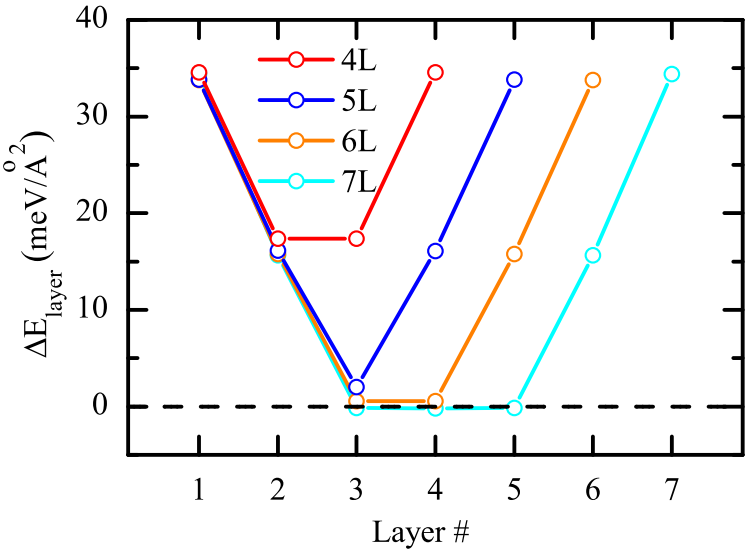

Figure 6.5: Atomic energy distribution on Au(100) slab. The atomic integration errors are smaller than 1 meV/Å^2. The bulk-like behavior of the center layer(s) of 5-, 6-, and 7-layer slabs indicates the sufficient thickness of slab calculation.

We represent the charge density and compute energy densities on a grid increasing from $60\times60\times350$ to $60\times60\times460$ for 4- to 7-layer supercells. Figure 6.5 shows the energy change from bulk value for each layer of atoms. Energy density integration of left (or right) 2 layers of 4-layer slab gives a surface energy of 52 ± 1 meV/Å^2; 5-, 6-, and 7-layer slabs all give a surface energy of 50 ± 1 meV/Å^2. Although the calculated surface energy of 4-layer surface energy is close to the 5-, 6- and 7-layer's values, the atomic energy distribution shows the

center layers of 4-layer slab have not reached the bulk-like behavior.

6.3 Au/TiO_2 interfaces

In traditional total energy calculation, the work of adhesion of forming an interface from two individual surfaces can be determined as:

$$E_{\text{adh}} = \frac{1}{A}(E_{\text{Au}} + E_{\text{TiO}_2} - E_{\text{Au/TiO}_2}), \tag{6.1}$$

where E_{Au} and E_{TiO_2} are the energy of relaxed Au surface and relaxed bare-support TiO_2 surface, respectively, calculated in the same supercell as the interface system, and $E_{\text{Au/TiO}_2}$ is the energy of the interface system.

Figure 6.6 shows the four different configurations of rutile TiO_2(110) substrates we consider. We start with a stoichiometric surface, and then reduce the surface by removing all bridging O atoms; both are 1×1 reconstructions. Pang *et al.* proposed an added-row 1×2 reconstruction for the rutile (110) surface, where one row of Ti atom with its sub-bridging O row are removed per 1×2 cell for a fully reduced surfaces. [79] Finally, removing the twofold coordinated O atoms from the added-row reconstruction gives a TiO reconstruction. While this reconstruction is not the lowest in energy, it provides the most stable Au/TiO_2 interface that also matches the experimentally observed geometry.

We calculate atomic energies by integrating the local energy density over gauge independent integration volumes. The changes in atomic energy helps understand the stability of formed interfaces. For the different Au/TiO_2 interfaces, the supercell configuration applied in the calculations contains six layers of Au periodically repeated in both directions parallel to the surface, eight trilayers of TiO_2 similarly periodic, and 10.5 Å vacuum region. Atomic relaxation is allowed only for six layer Au atoms and three interfacial layers of TiO_2. The equilibrium positions of the atoms are determined by requiring the force on each relaxed

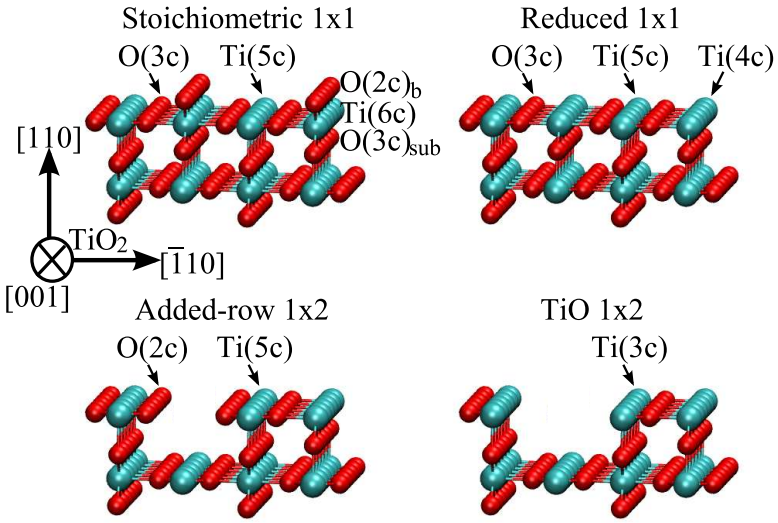

Figure 6.6: Geometry for four different $TiO_2(110)$ surface structures. Upper two are stoichiometric 1×1, and reduced 1×1; bottom two are added-row 1×2 reconstruction, and TiO 1×2 reconstruction. The stoichiometric structure has bridging O $(2c)_b$ atoms above the flat Ti (5c) / (6c) and O (3c) plane. Removal of the bridging O produces a reduced surface, with four- and five-fold coordinated Ti. The added-row reconstruction removes every other row of Ti (4c) atoms with subsurface bridging O $(3c)_{sub}$ for a 1×2 reconstruction, with two-fold coordinated O. Finally, additional reduction of the added-row reconstruction, by removing the O (2c) atoms neighboring the removed row, produces the TiO reconstruction with three-fold coordinated Ti.

atom to be smaller than 0.02 eV/Å.

6.3.1 Au(111)//TiO_2(110) 1×1: Stoichiometric and reduced interfaces

Both interfacial reconstructions on 1×1 surfaces use a similar geometry for relaxation. Along the direction Au[$1\bar{1}0$]//TiO_2[001], a single repeat length of Au and TiO_2 gives a 1% lattice mismatch. This agrees with STEM measurements showing registry even up to 10 layers from the interface. Along the direction Au[$\bar{1}\bar{1}2$]//TiO_2[$\bar{1}10$], a repeat length of 4 for Au matches with a repeat length of 3 for TiO_2, producing a total 3.6% lattice mismatch strain at the interface. The supercell is 2985 Å^3 in volume, containing 48 Au, 48 Ti, and 96 O atoms in the interface configuration with stoichiometric TiO_2 surface, and 3 fewer O atoms for the

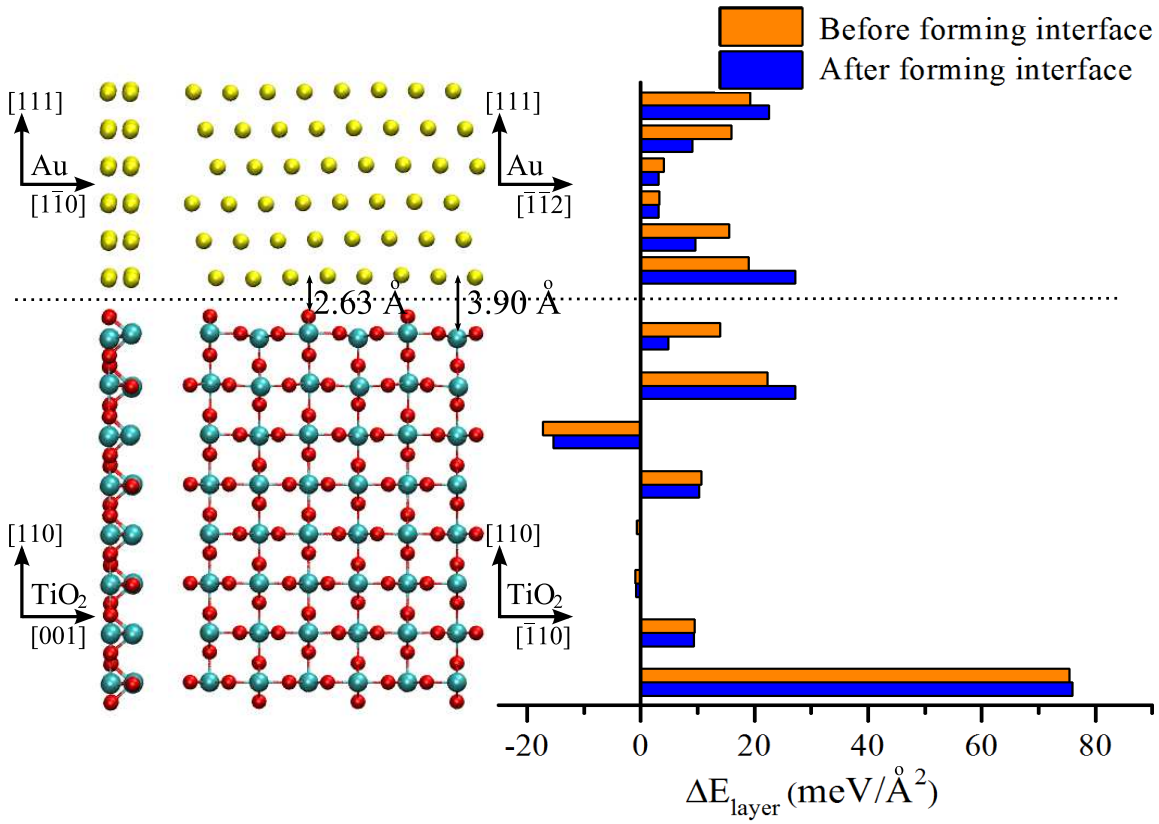

Figure 6.7: Geometry and energy of Au(111) on the stoichiometric TiO_2(110) surface. All six Au(111) planes and the top three trilayers of the eight trilayer TiO_2(110) surface are relaxed. The atomic energy on each layer is referenced to the bulk value, and shown before (orange) and after (blue) forming the interface. The interfacial distance relaxes to 3.90 Å between Au and Ti layers, and 2.63 Å between Au and bridging-O layers.

reduced TiO_2 surface. In each case, we relax all six layers of Au(111) and the top three layers of TiO_2(110); this allows us to determine the interlayer spacing at the interface, and by using energy density calculations, we can ignore any spurious energy changes due to the far Au surface.

Figure 6.7 and Figure 6.8 show the geometry of the relaxed Au(111) on stoichiometric and reduced TiO_2(110) reconstructions. The interfacial distance between Au and Ti layers relaxed to 3.90 Å with stoichiometric TiO_2 surface, and 2.79 Å in the configuration with reduced TiO_2 surface. The interfacial distance between Au and bridging-O layers relaxed to 2.63 Å in the configuration with stoichiometric TiO_2 surface, and between Au and in-plane O

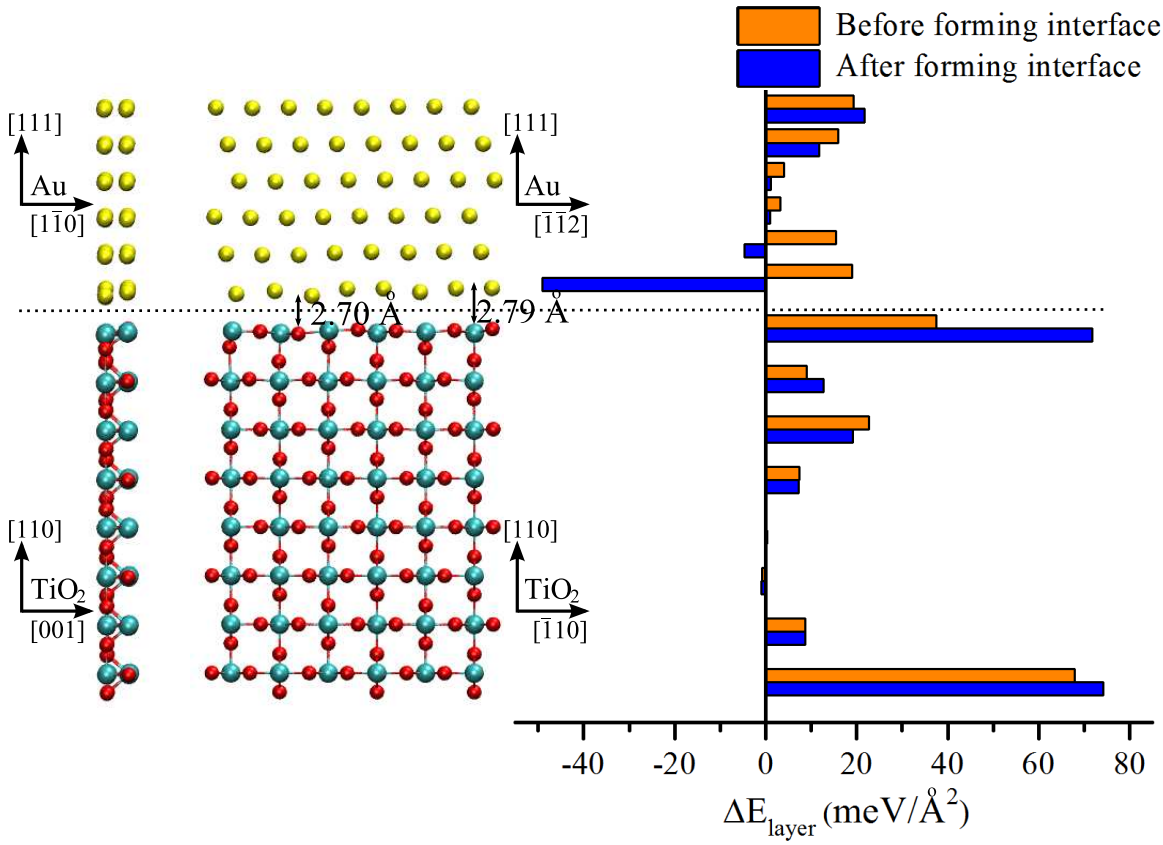

Figure 6.8: Geometry and energy of Au(111) on the reduced TiO_2(110) surface. Energy per layer in the reference of bulk value is given before (orange) and after (blue) forming the interface. The interfacial distance relaxes to 2.79 Å between Au and Ti layers; and 2.70 Å between Au and in-plane O layers. The geometry reduces the energy of the surface Au layer to a more stable configuration than the stoichiometric TiO_2 surface.

layers relaxed to 2.70 Å in the configuration with reduced TiO_2 surface. From total energy, the work of adhesion of the interface with stoichiometric TiO_2 surface is 7 meV/Å^2, while the work of adhesion of the interface with the reduced TiO_2 surface is 54 meV/Å^2. The differences in interlayer spacing and energy is due to the presence or absence of bridging oxygen atoms on the TiO_2 surface. Energy density shows that TiO_2 layers reach bulk behavior by the fifth layer from interfaces. We integrate the energy density over two Au layers and four TiO_2 layers to evaluate the work of adhesion strictly from changes in energy near the interface. This gives a work of adhesion of 4 ± 2 meV/Å^2 to the stoichiometric TiO_2 surface, and 53 ± 1 meV/Å^2 to the reduced TiO_2 surface. The work of adhesion is

primarily due to a decrease in energy of the Au surface layer at the reduced TiO_2 surface. This suggests that the main effect of removing bridging oxygen is to provide a flat surface for Au(111) layers to adhere, and that the TiO_2 surface energy change is significantly less than the Au surface change.

6.3.2 Au(100)//TiO_2(110) 1×1: Stoichiometric and reduced interfaces

Both interfacial reconstructions on 1×1 surfaces use a similar geometry for relaxation. Along the direction Au[01$\bar{1}$]//TiO_2[001], a single repeat length of Au and TiO_2 gives a lattice mismatch 1%. Along the direction Au[011]//TiO_2[$\bar{1}$10], a repeat length of 9 for Au matches with a repeat length of 4 for TiO_2, producing a lattice mismatch 0.9%. This supercell is 3854 $Å^3$ in volume, containing 54 Au, 64 Ti, and 128 O atoms for stoichiometric case, and 4 fewer O atoms for reduced case. As before, we relax all six layers of Au(100) and the top three layers of TiO_2(110); this allows us to determine the interlayer spacing at the interface, and by using energy density calculations, we can ignore any spurious energy changes due to the far Au surface.

Figure 6.9 and Figure 6.10 show the geometry of the relaxed Au(100) on stoichiometric and reduced TiO_2(110) reconstructions. The interfacial distance between Au and Ti layers relaxed to 3.63 Å with stoichiometric TiO_2 surface, and 2.64 Å in the configuration with reduced TiO_2 surface. The interfacial distance between Au and bridging-O layers relaxed to 2.33 Å in the configuration with stoichiometric TiO_2 surface, and between Au and in-plane O layers relaxed to 2.55 Å in the configuration with reduced TiO_2 surface. From total energy, the work of adhesion is 3 meV/$Å^2$ of the interface with stoichiometric TiO_2, while the work of adhesion of the interface with the reduced TiO_2 surface is 55 meV/$Å^2$. The differences in interlayer spacing and energy is due to the presence or absence of bridging oxygen atoms on the TiO_2 surface. Energy density shows that TiO_2 layers reach bulk

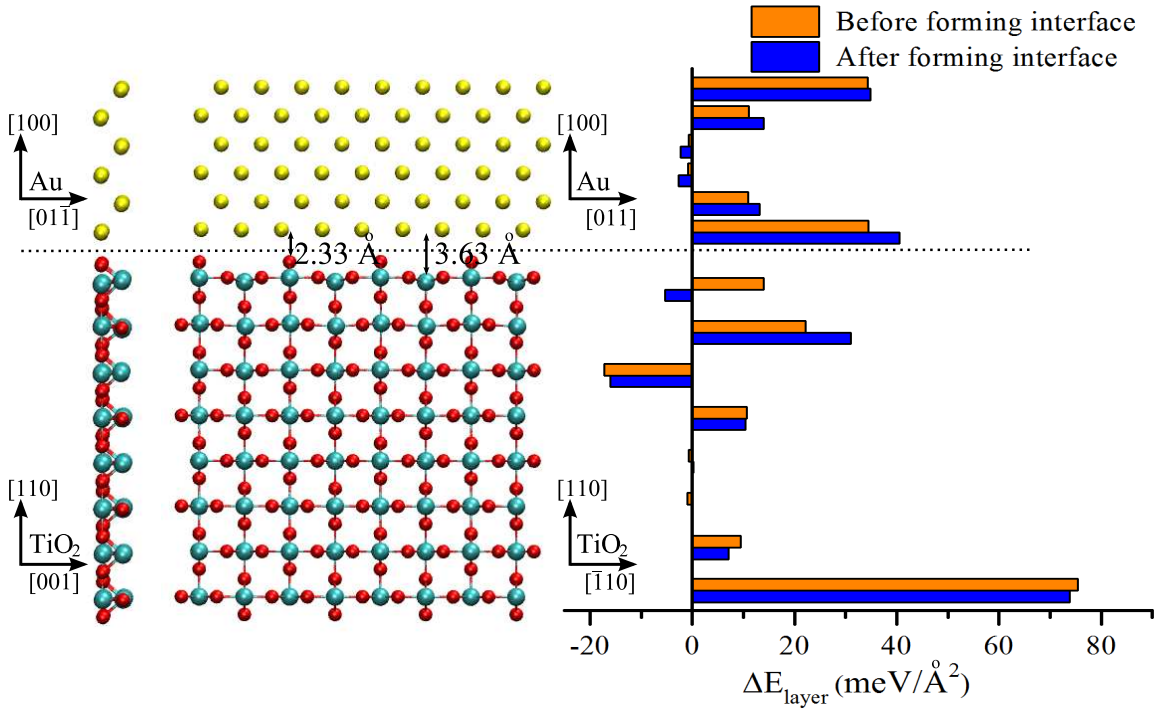

Figure 6.9: Geometry and energy of Au(100) on the stoichiometric TiO_2(110) surface. All six Au(100) planes and the top three trilayers of the eight trilayer TiO_2(110) surface are relaxed. The atomic energy on each layer is referenced to the bulk value, and shown before (orange) and after (blue) forming the interface. The interfacial distance is 3.63 Å between Au and Ti layers, and 2.33 Å between Au and bridging O layers.

behavior by the fifth layer from interfaces. We integrate the energy density over two Au layers and four TiO_2 layers to evaluate the work of adhesion strictly from changes in energy near the interface. This gives a work of adhesion of 1 ± 1 meV/Å^2 to the stoichiometric TiO_2 surface, and 64 ± 1 meV/Å^2 to the reduced TiO_2 surface. Similar to Au(111)//TiO_2(110) reduced interface, atomic energy at interface decreases in Au layer, and increases in TiO_2 layer in the reduced case during forming the interface causes this structure stabler than stoichiometric case. The energy of TiO_2 free surfaces experience spurious energy changes from the interface formation. Therefore, the integration of energy density over interfacial region reduces the finite-size error, and provides more accurate work of adhesion or interfacial energy.

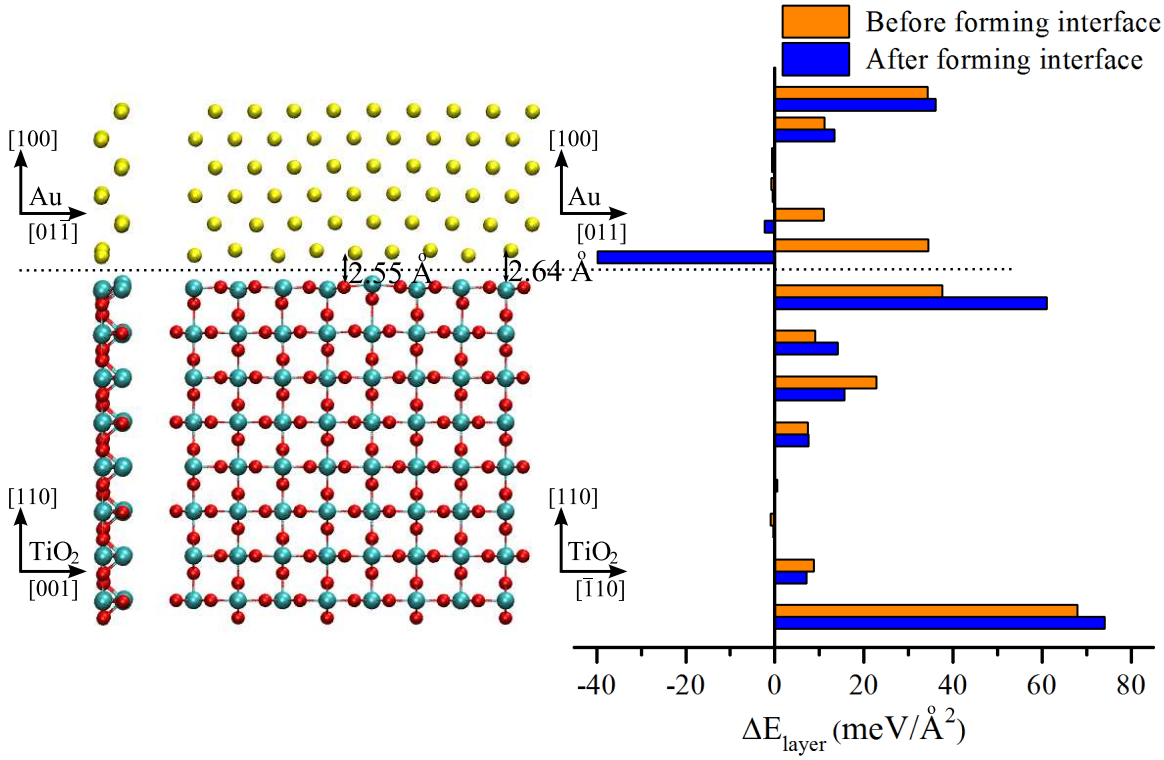

Figure 6.10: Geometry and energy of Au(100) on the reduced TiO_2(110) surface. The atomic energy on each layer is referenced to the bulk value, and shown before (orange) and after (blue) forming the interface. The interfacial distance is 2.64 Å between Au and Ti layers, and 2.55 Å between Au and in-plane O layers. The Au layer energy is reduced while the TiO_2 layer energy increases for a net work of adhesion of $64 \pm 1\,\mathrm{meV/Å^2}$.

6.3.3 Au(111)//TiO_2(110) 1×2: Added-row and TiO reconstructions

Both interfacial reconstructions on 1×2 surfaces use a similar geometry for relaxation. Along the direction Au[$1\bar{1}0$]//TiO_2[001], a single repeat length of Au and TiO_2 gives a 1% lattice mismatch as for the 1×1 reconstructions. Along the direction Au[$\bar{1}\bar{1}2$]//TiO_2[$\bar{1}10$], a repeat length of 5 for Au matches with a repeat length of 4 for TiO_2, producing a total 2.9% lattice mismatch strain at the interface; the different periodicity is required for a 1×2 reconstruction. The supercells contains 62 Au, 62 Ti, and 122 O atoms in the interface configuration with added-row TiO_2 reconstruction, and 4 fewer O atoms for the TiO reconstruction. In each

case, we relax all six layers of Au(111) and the top three layers of TiO_2(110); this allows us to determine the interlayer spacing at the interface, and by using energy density calculations, we can ignore any spurious energy changes due to the far Au surface.

Added-row reconstruction

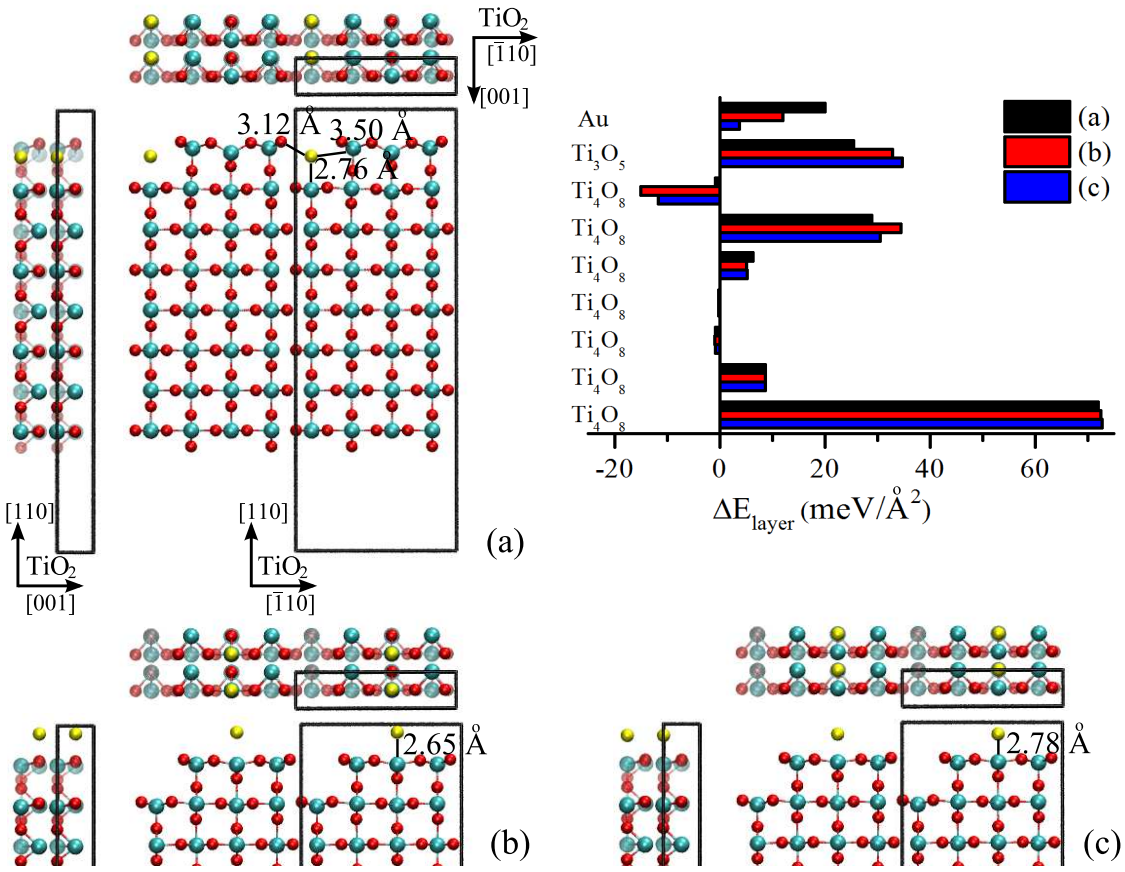

Figure 6.11: Three different configurations of a single Au row on the TiO_2 added-row reconstruction: (a) in the missing Ti row with 4 nearest neighboring O atoms; (b) on top of the TiO_2 surface directly above a Ti atom; (c) on top of the TiO_2 surface bridging between two Ti atoms. Au atoms are in gold, and the wireframe shows the supercell. Opaque atoms are on the top layer while transparent atoms are on lower layers. The (c) configuration has lowest total energy, 6.3 meV/Å^2 lower than (b) configuration, and 15.7 meV/Å^2 lower than (a) configuration. From the energy calculations, the Au row controls the total energy, with the largest increase in energy occurring from filling the missing Ti-O row in the surface layer; hence, we expect to see a mixing of the TiO_2 surface with Au only *after* coverage by a gold nanoparticle.

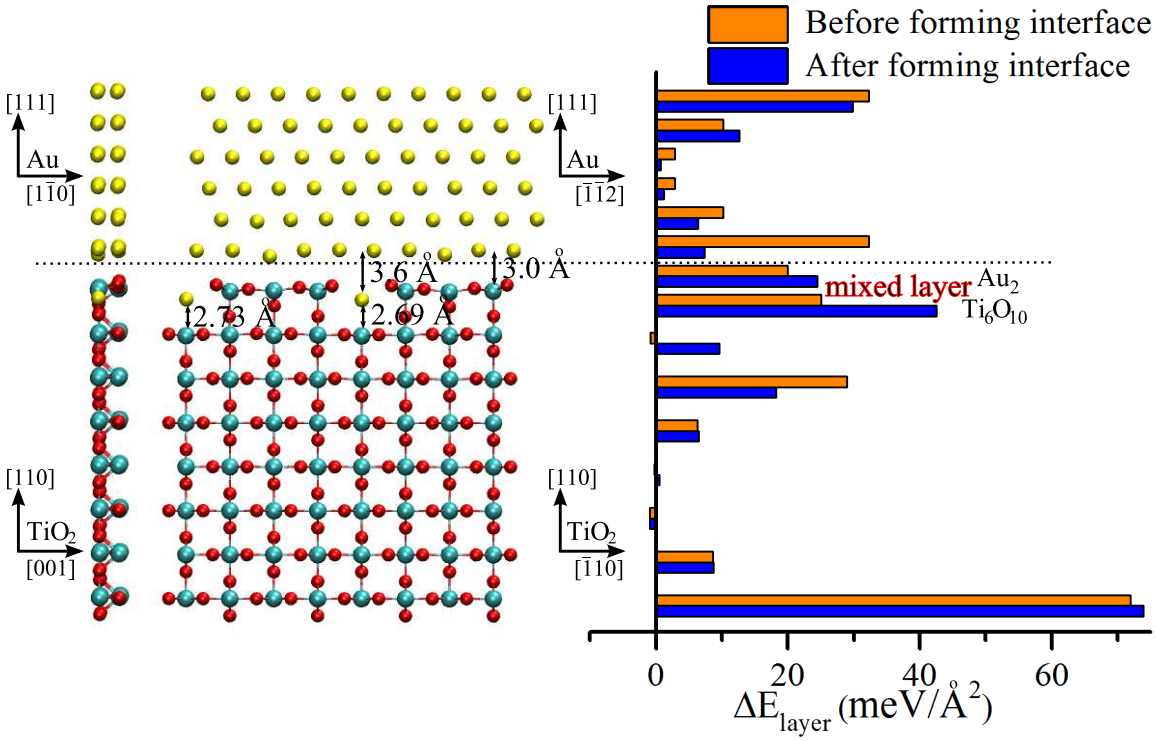

Figure 6.12: Geometry and energy of the Au(111)//TiO_2 added-row reconstruction. Atomic energy per layer in the reference of bulk value is given before (orange) and after (blue) forming the interface. The interfacial distance between Au layer with mixed layer is about 3.4 Å, and the work of adhesion is 7 meV/Å^2. While the Au surface layer reduces its energy, the TiO_2 layer increases in energy as the oxygen atoms that neighbor the in-surface Au rows are unable to relax out of the (110) plane; hence, the Ti_6O_{10} layer increases in energy.

The added-row reconstruction for the 1×2 rutile (110) surface removes one row of Ti atom with its sub-bridging O row per 1×2 cell for a fully reduced surfaces. [79] Experimental observations of the interface find a mixed TiO_2-Au layer with 1×2 periodicity; to build our interface and compute the work of adhesion, we consider different configurations to attaching a row of Au atoms on added-row reconstruction in Figure 6.11. After geometry relaxation, the configuration of each Au atom sitting on the top of two Ti atoms with 4 neighboring O atoms is the most stable; there is an energy cost of 15.7 meV/Å^2 to place a Au row into the missing row of TiO_2; again, the energy density shows that the energy of Au dominates the stability.

Figure 6.12 shows the geometry of the relaxed Au(111) on added-row TiO_2 reconstruction.

The interfacial distance between Au and the mixed interfacial layer is 3.4 Å. This larger distance is due to the displacement of oxygen atoms neighboring the interfacial Au rows. From total energy, the work of adhesion of the interface is $7\,\mathrm{meV/Å^2}$. We integrate the energy density over two Au interfacial layers, one mixed interfacial layer and three next TiO_2 interfacial layers and subtract the corresponding energy density integration in Au layers and the ground-state configuration of an Au row on TiO_2, Figure 6.11(a). This energy density calculation gives a work of adhesion of $6 \pm 1\,\mathrm{meV/Å^2}$. After forming interface, the atomic energy of Au interfacial layer drops, while the atomic energy of TiO_2 in the mixed layer increases. The increase in the TiO_2 layer, made up of Ti_3O_5 layers, is due to the constraint placed on the oxygen atoms neighboring to the intermixed Au row in the mixed layer.

TiO reconstruction

The added-row reconstruction can be further reduced by removing the twofold coordinated O atoms on the TiO_2 surface layer to form a new TiO 1×2 reconstruction. This reconstruction is suggested by the energy density calculations above as a possible route to increase the work of adhesion. We build our interface in a similar manner as for the added row reconstruction, and consider different configurations to attach one row of Au atoms on the reconstruction in Figure 6.13. After geometry relaxation, both the Au row in the missing row of Ti and on the surface have large, but similar, energies (a difference of $0.8\,\mathrm{meV/Å^2}$). The increase in surface energy is entirely due to the first TiO_2 layer, suggesting that further reduction to TiO is unfavorable *without* an interfacial layer of gold to "protect" the surface.

Figure 6.14 shows the geometry of the relaxed Au(111)//TiO reconstruction interface. Despite the higher energy of the TiO reconstruction, it produces a *stable* interface configuration with Au(111). The interfacial distance between the Au layer and mixed interfacial layer is 2.44–2.45 Å; the closer attachment distance compared with the added-row reconstruction is due to the removed oxygen atoms in the interfacial layer. From total energy, the work of adhesion of the interface is $99\,\mathrm{meV/Å^2}$. We integrate the energy density over two Au inter-

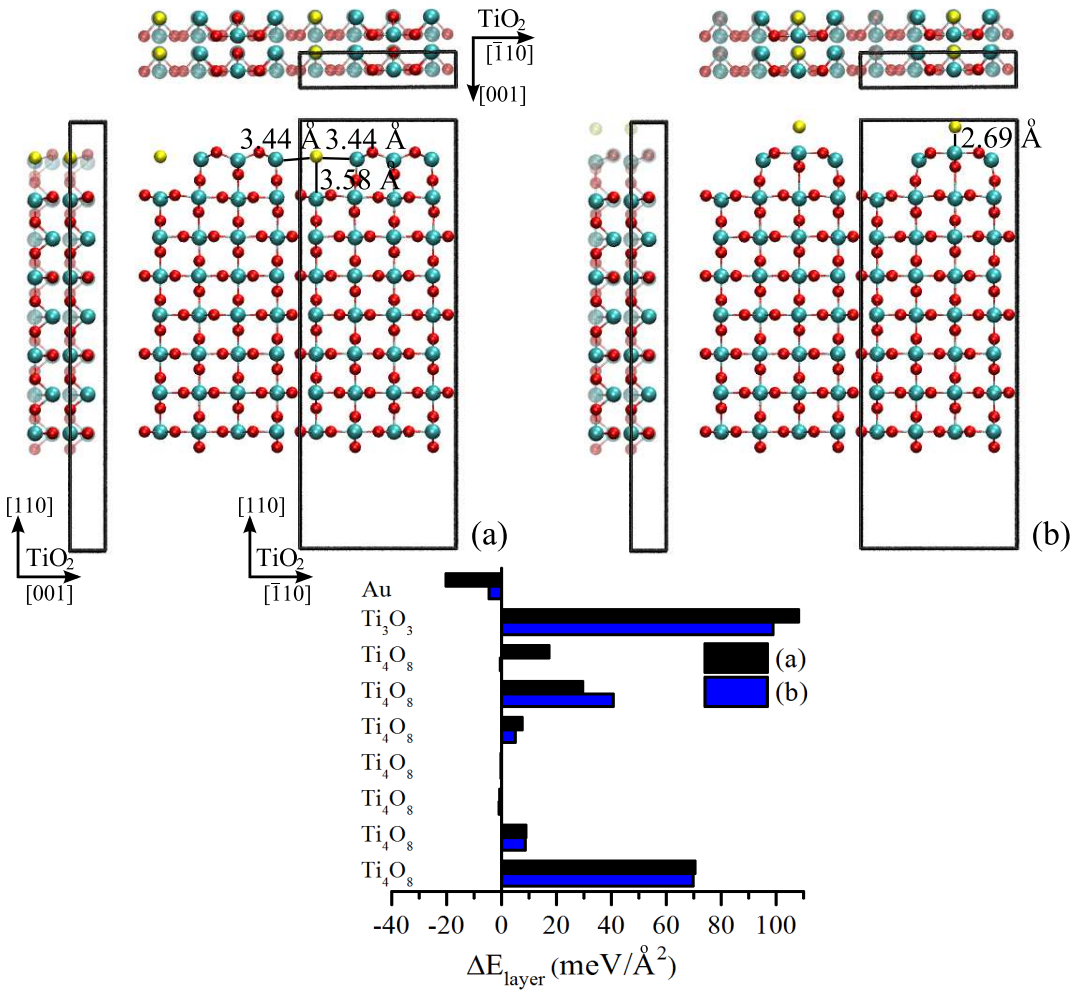

Figure 6.13: Two different configurations of a single Au row on the TiO reconstruction: (a) in the missing Ti row with 4 nearest neighboring O atoms; (b) on top of the TiO_2 surface bridging between two Ti atoms. Au atoms are in gold, and the wireframe shows the supercell. Opaque atoms are on the top layer while transparent atoms are on lower layers. The energy of the (b) configuration is $0.8\,\mathrm{meV/Å^2}$ lower than (a) configuration. Adding Au into the missing row only slightly increases the energy of the Au row; this increase is much less than for the TiO_2 added-row reconstruction. However, the TiO reconstruction is a higher energy surface than the added-row reconstruction.

facial layers, one mixed interfacial layer and three next TiO_2 interfacial layers and subtract the corresponding energy density integration in Au layers and the ground-state configuration of an Au row on TiO Figure 6.13(a). This energy density calculation gives a work of adhe-

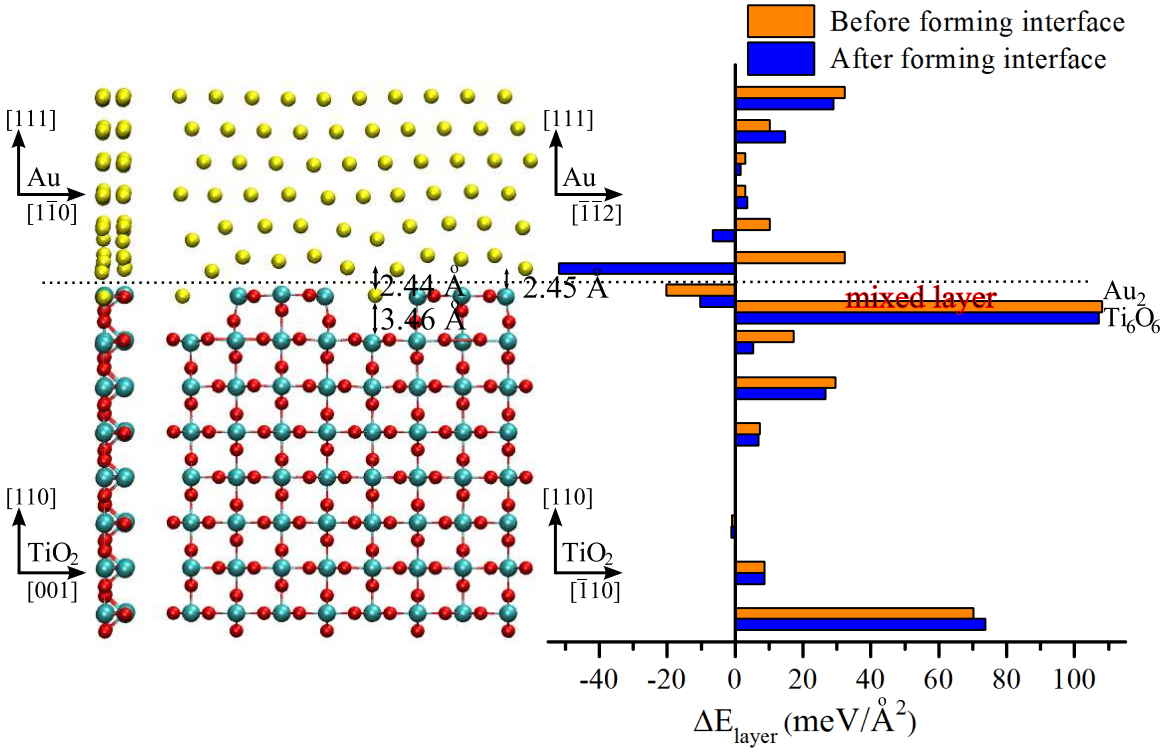

Figure 6.14: Geometry and energy of the Au(111)//TiO reconstruction. Atomic energy per layer in the reference of bulk value is given before (orange) and after (blue) forming the interface. The interfacial distance between Au layer with mixed layer is 2.45 Å, compared with experimental observation of 2.5 Å. The work of adhesion is 107 meV/Å^2 from energy density integration compared with the Au(111) and TiO reconstruction filled with a Au row. The stability of the interface comes from a reduction in the Au surface energy with no penalty in the mixed layer, as occurs with the added-row reconstruction.

sion of 107 ± 1 meV/Å^2; the difference with the total energy calculation is due to spurious changes in the free TiO_2 surface that the energy density calculation removes. Similar to the Au(100)//reduced TiO_2(110) interface, we observe a remarkable drop of atomic energy on Au interfacial layer. In addition, the mixed layer energy sees only a small change leading to a stabilized interface. To compute the true work of adhesion, however, we must account for the energy change due to a further reduction from the added-row reconstruction to the TiO reconstruction.

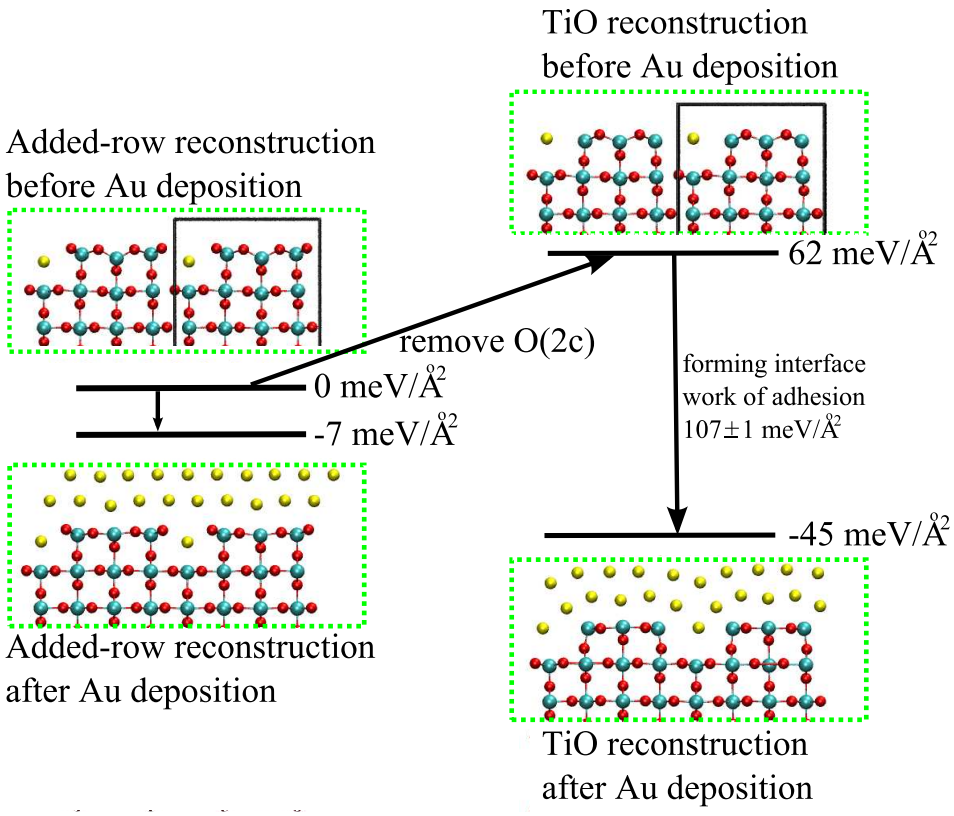

Figure 6.15: Evaluation of work of adhesion for Au(111) on 1×2 TiO_2(110) reconstructions. The TiO reconstruction leads to a stable interface after Au deposition as the energy required to remove additional oxygen atoms from the added-row reconstruction is offset by a larger reduction in energy when forming the interface. This is an interesting example of a interfacial reconstruction that is stabilized solely in the presence of the interface. Compared with the other simple added-row reconstruction which produces a small work of adhesion due to distortions in the mixed layer, the new TiO interfacial reconstruction explains the observed 1×2 reconstruction, the interlayer spacing, and is energetically favorable.

Work of adhesion

Figure 6.3.3 shows the relative energies for the different configurations to produce the two different 1×2 reconstructions of Au(111)//TiO_2(110). Au(111) adhered to the new TiO reconstruction is the most stable interface configuration with work of adhesion of 107 meV/Å^2 and an interfacial distance 2.45 Å agreeing with the STEM experimental observation 2.35 ± 0.16 Å; however, that energy change occurs relative to the higher energy TiO surface. The relative difference between the added-row reconstruction and the TiO reconstruction means

that a single Au row on the TiO reconstruction is less stable by $62\,\mathrm{meV/\AA^2}$; hence, the TiO reconstruction produces a stable configuration with work of adhesion of $45\,\mathrm{meV/\AA^2}$ after Au deposition. This is lower than simply adhering to the added-row reconstruction, which has a work of adhesion of $7\,\mathrm{meV/\AA^2}$. It should be noted that the intermediate configuration of TiO without Au(111) is unstable, and is needed to compute relative energies; given the higher surface energy, it is unlikely that further oxygen reduction occurs before the growth of Au(111) layers.

6.4 Summary

		E_{adh} (meV/Å^2)	$d_{\text{Au-Ti}}$ (Å)
Stoichiometric	1×1	$7 \mid 4\pm1$	3.90
Reduced	1×1	$54 \mid 53\pm1$	2.79
Added-row	1×2	$7 \mid 7\pm1$	3.00
TiO	1×2	$38 \mid 45\pm1$	2.45
Experiment [11]	1×2	28 ± 7	2.35 ± 0.16

Table 6.7: Comparison of different Au(111)//TiO_2(110) interfaces. The two different work of adhesions are from the total energy calculation of Eqn. (6.1), and the energy density integration; the latter accounts for finite-size errors in the supercell calculation. The Au(111) //TiO reconstruction agrees with experimental observation in three factors: the 1×2 symmetry, the work of adhesion W_{adh}, and the Au-Ti separation distance $d_{\text{Au-Ti}}$. The work of adhesion $45\pm1\,\mathrm{meV/\AA^2}$ is at the upper limit of $28\pm7\,\mathrm{meV/\AA^2}$ (see details in Section 1.1).

We apply density-functional theory energy density calculations to several interfacial configurations of Au/TiO_2 to determine the equilibrium structure that matches experimental measurements. Both Au(111) and (100) prefer attaching to reduced rutile TiO_2(110) surfaces over stoichiometric surfaces. Comparison of Au(111) attaching on two TiO_2(110) 1×2 reconstruction cells shows that our new TiO reconstruction leads to the most stable interface configuration with interfacial distance 2.45 Å, and net work of adhesion $45\,\mathrm{meV/\AA^2}$. Atomic energy variation during interface formation demonstrates that the attraction of top Au interfacial layer leads to a stable structure. The energy density computation also identifies

spurious changes to atomic energies on the free-surfaces during the formation of an interface, which affect the computation of work of adhesion from total energy calculations; these finite-size errors are removed. Besides the new reconstruction, where further reduction of the interface is possible when "protected" by an epitaxial gold layer, this demonstrates the power of energy density computation to guide the identification of stable defect structures.

Chapter 7

Conclusion

The work presented in this thesis is the development and application of a method to decompose the energy of a material into contributions from different spatial regions, so that a single calculation can provide much more information than the total energy. The generality of the method is demonstrated by application to various defects in crystals, and an extensive study is carried out to provide understanding of the stability and properties of Au/TiO_2 interfaces, chosen because of the relevance for gold nanoparticles supported on TiO_2 that are catalysts for oxidation of CO to CO_2.

The theoretical method is based upon a formulation of the energy density that is divided into 4 terms: a localized kinetic energy density; a total Coulomb energy density including all interactions among electrons and ions; the exchange-correlation energy density in density functional theory; and short-ranged terms localized in the core regions around individual atoms. As is well-known the kinetic energy at any spatial point is not unique; however, one can define a kinetic energy density that is unique except for terms proportional to the Laplacian of electron density that can be thought of as a gauge transformation. Nevertheless, unique gauge independent kinetic energies can be obtained as integrals over any volume bounded by a zero-flux surface of the gradient of the electron density. This aspect of the work is the same as defined by Bader [17], but the other terms are treated differently. We treat the Coulomb energy in a way analogous to the kinetic energy using the fact that Coulomb energy density can be expressed in two forms in terms of the potential and charge or in terms of the electric field squared (the Maxwell expression). Thus it is uniquely defined in a volume bounded by zero-flux surface of electrostatic field, i.e., a zero charge volume that

is different from the Bader volume. The exchange-correlation energy can be expressed in various functionals, all are given as a function of electron density and its gradients. The short-ranged terms in the core regions are treated differently in different methods; we have derived the appropriate forms for pseudopotential and all-electron PAW methods, which lead to terms assigned to each atom and not explicitly given as densities.

Another crucial part in this work is deriving efficient, accurate methods to carry out the integration of energy density. A novel weight method is proposed to calculate energy density integration bounded by zero-flux surface of electron density gradient or potential gradient, and generally is applied on constructing zero-flux surface of the gradient of an arbitrary functional. Start from the local density maxima, grid points are sorted in density descending order. Grid points can then be fractionally weighted one by one following the atomic weights of its neighbors with larger density. This method depends upon the formulation of flow across that dividing surfaces between the cells of two neighboring grid points that can be applied to uniform or non-uniform grids. The algorithm is robust, efficient with $O(N)$ computing time, and more accurate than other existing grid-based algorithm.

The next step is the physical meaning of the energy density integrated over the zero-flux volumes. In many cases, it can be shown that the integrals can be used to calculate different physically meaningful energies from one calculation. An example is an interface energy, with the added benefit that the variation of the energy density shows explicitly the extent of the interfacial regions and the regions that can be considered to be bulk. In addition, the zero-flux volumes always surround each atom defining volumes that can be assigned to each atom, and one can associate the energy of the volumes with the atom. This provides a physical picture of the energy density on an atomic scale that can be used to interpret the regions responsible for binding at the interface.

Various tests and applications have been studied from surface energies to point defect energies; from semiconductors to metals, metal oxides to demonstrate the ability of our method to isolate defect energies in a DFT calculation. The traditional total energy method

is also applied to calculate the defect energies, and is in agreement within error bars with the results from current energy density method. Furthermore, the current energy density method has advantages in several aspects. First, in calculating a defect or surface energy, the total energy calculation method can calculate only one energy as the difference between two calculations. In contrast, from the energy density one can calculate energies integrated over various regions at once. For example, we have shown that the relaxed and the un-relaxed surface energies of $TiO_2(110)$ can be obtained from only one slab calculation. Second, unlike the convergence test of the traditional total energy calculation, the required thickness of a slab can be directly determined by observing the variation of atomic energy from the bulk value. Third, the energy density calculation reveals some properties can not be obtained by total energy calculation. For example, the fifth shell confinement on Si monovacancy is observed from energy density perturbation around the point defect; the O interstitial in the hexagonal close-packed (HCP) Ti crystal demonstrates a Friedel oscillation both in charge density and energy density.

We especially apply the energy density method to studying the interface of Au/TiO_2 for its importance as a model oxidation catalyst. DFT calculations are performed on several different Au/TiO_2 interfaces including: Au(111) and Au(100) on the top of $TiO_2(110)$ stoichiometric and reduced surfaces; Au(111) on added-row 1×2 $TiO_2(110)$ reconstruction and a new proposed TiO 1×2 reconstruction. Both Au(111) and (100) prefer attaching to reduced rutile $TiO_2(110)$ surfaces over stoichiometric surfaces. The DFT and energy density calculations with geometry relaxation show that Au(111) on our new TiO reconstruction is the most stable interface with the net work of adhesion of 45 ± 1 standing on the uplimit of $28 \pm 7\,\mathrm{meV/\AA^2}$ and an interfacial distance $2.45\,\mathrm{\AA}$ agreeing with the STEM experimental observation $2.35 \pm 0.16\,\mathrm{\AA}$. Atomic energy variation during forming interfaces demonstrates that the attraction of top Au interfacial layer leads to a stable structure. In the cases that the atomic energies on the back layers of Au and TiO_2 change during forming interface, energy density integration gives more accurate interfacial energy than total energy calculation.

Appendix A

Atomic Units

Hartree atomic units (abbreviated "a.u.") are chosen to describe the electron properties in atoms, particularly for the Bohr model of hydrogen atom in its ground state.

	Symbol	Atomic units	Conventional units
Electron mass	m_e	1	$9.10938215(45) \times 10^{-31}\,\mathrm{kg}$
Electron charge	e	1	$1.602176487(40) \times 10^{-19}\,\mathrm{C}$
Reduced Planck's constant	$\hbar = h/(2\pi)$	1	$1.054571628(53) \times 10^{-34}\,\mathrm{J}$
			$= 6.58211899(16) \times 10^{-16}\,\mathrm{eV\,s}$
Electric constant	$1/(4\pi\varepsilon_0)$	1	$8.9875517873681 \times 10^{9}\,\mathrm{kg\,m^3\,s^{-2}\,C^{-2}}$
Bohr radius	$a_0 = 4\pi\varepsilon_0\hbar^2/(m_\mathrm{e}e^2)$	1	$5.2917720859(36) \times 10^{-11}\,\mathrm{m}$
Hartree energy	$E_\mathrm{h} = m_\mathrm{e}e^4/(4\pi\varepsilon_0\hbar)^2$	1	$4.35974394(22) \times 10^{-18}\,\mathrm{J}$
			$= 27.21138386(68)\,\mathrm{eV}$
Electron volt	eV		$1.602176487(40) \times 10^{-19}\,\mathrm{J}$
Velocity	$a_0E_\mathrm{h}/\hbar$	1	$2.1876912541(15) \times 10^{6}\,\mathrm{m\,s^{-1}}$
Speed of light	c		$299792458\,\mathrm{m\,s^{-1}}$
Boltzmann constant	k_B	1	$1.3806504(24) \times 10^{-23}\,\mathrm{J\,K^{-1}}$
			$= 8.617343(15) \times 10^{-5}\,\mathrm{eV\,K^{-1}}$

Table A.1: Hartree atomic units. CODATA Internationally recommended values of the Fundamental Physical Constants (http://physics.nist.gov/cuu/Constants/index.html).

Appendix B

Electron localization function

In physics, the electron localization function (ELF) is a method of measuing spatial localization of the reference electron in multielectronic systems, originally defined by Becke and Edgecombe in 1990 [87].

Before talking about the ELF, we divide the kinetic energy density into two parts [41],

$$t(\mathbf{r}) \;=\; t_\rho(\mathbf{r}) + t_{\mathrm{x}}(\mathbf{r})\,. \tag{B.1}$$

The former is the dependence of density $\rho(\mathbf{r})$; the later is the exchange part of kinetic energy density. In section 3.1, we show that the gauge dependent term in kinetic energy density is $-\frac{1}{4}\nabla^2\rho(\mathbf{r})$. The density dependent part of asymmetric and symmetric kinetic energy density turns out to be

$$\begin{aligned} t_\rho^{(a)} \;&=\; -\frac{1}{2}\rho^{\frac{1}{2}}\nabla^2\rho^{\frac{1}{2}} = -\frac{1}{4}\nabla^2\rho + \frac{1}{8}\frac{(\nabla\rho)^2}{\rho}\,, \\ t_\rho^{(s)} \;&=\; \frac{1}{2}|\nabla\rho^{\frac{1}{2}}|^2 = \frac{1}{8}\frac{(\nabla\rho)^2}{\rho}\,. \end{aligned} \tag{B.2}$$

The difference of $t_\rho^{(a)}$ and $t_\rho^{(s)}$ is also equal to $-\frac{1}{4}\nabla^2\rho(\mathbf{r})$. Therefore, we conclude that the exchange kinetic energy density is well-defined, and the non-uniqueness of the kinetic energy density is only due to the gauge dependence of density dependent term.

The ELF for each spin σ is defined by Beche and Edgecombe as

$$\mathrm{ELF}(\mathbf{r}) \;=\; \frac{1}{1+(\chi^\sigma(\mathbf{r}))^2}\,, \tag{B.3}$$

where $\chi^\sigma(\mathbf{r}) = t^\sigma_{\mathrm{X}}/t^\sigma_{\mathrm{TF}}$, with $t^\sigma_{\mathrm{X}} = \frac{1}{2}\sum^\sigma_{n\mathbf{k}} f^\sigma_{n\mathbf{k}}|\nabla\psi^\sigma_{n\mathbf{k}}|^2 - \frac{1}{8}\frac{(\nabla\rho^\sigma)^2}{\rho^\sigma}$ the well-defined exchange kinetic energy density and $t^\sigma_{\mathrm{TF}} = \frac{3}{10}(6\pi^2)^{2/3}(\rho^\sigma)^{5/3}$ the Thomas-Fermi expression in homogeneous electron gas. $\chi^\sigma(\mathbf{r})$ is a dimensionless quantity that expresses electron localization with respect to the homogeneous electron gas. The value is in the range $0 \leq \mathrm{ELF}(\mathbf{r}) \leq 1$, whose upper limit $\mathrm{ELF}(\mathbf{r}) = 1$ corresponding to perfect localization and the value $\mathrm{ELF}(\mathbf{r}) = 1/2$ corresponding to homogeneous gas-like behavior. The well-known application of ELF is visualization of atomic shell structure, whereas the analysis of the radial density distribution $4\pi r^2\rho(r)$ or the Laplacian of the density $\nabla^2\rho$ fails to reveal more than five atomic shell structures for heavier atoms.

Appendix C

Ewald Sum

The Ewald sum named after P. P. Ewald is a method to calculate the electrostatic energy of periodic system. Since the Coulomb energy contains short range and long range interactions, the idea of the Ewald sum is to decompose the electrostatic potetial into short range and long range components. Keep the short range component of real space sum in real space, and Convert the long range component of real space sum to Fourier space. The later is equivalent to the short range component of Fourier space sum. So that both terms converge fast, and the truncation only costs little loss of accuracy.

For an electrically neutral system including N-particles with positive and negative point charges, $\sum_{i=1}^{N} q_i \delta(\mathbf{r} - \mathbf{r}_i) = 0$, under periodic boundary condition, the electrostatic energy is given by

$$\begin{aligned} E &= \frac{1}{2} \sum_{i,j}^{N} \sum_{\mathbf{n}} \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{nL}|} \\ &= \frac{1}{2} \sum_{i=1}^{N} q_i V(\mathbf{r}_i) \,, \end{aligned} \tag{C.1}$$

where q_i and $\mathbf{r}_i$ are the charge and the position of the i^{th} particle in the unit cell and $\mathbf{nL}$ is a set of supercell translation vectors. The electrostatic potential at particale i is $V(\mathbf{r}_i) = \sum'_{j,n} \frac{q_j}{|\mathbf{r}_{ij}+\mathbf{nL}|}$ with the prime indicates the exclusion of the self interaction, that particale i can interact with all other particles and their images, can interact with all its images, but not, with itself.

One typical way to separate the short range and long range components is to add a

ficticious charge distribution of opposite sign $-\rho(\mathbf{r})$. Thus the electrostatic potential consists of two parts: the short-ranged component is due to the screened charge $q_i\delta(\mathbf{r}-\mathbf{r}_i)-\rho(\mathbf{r})$; the long-ranged component is due to the compensating charge $\rho(\mathbf{r})$. The ficticious charge distribution is usually chosen as a Gaussian function

$$\rho(\mathbf{r}) \;=\; \sum_{j=1}^{N}\sum_{\mathbf{n}} q_j \left(\frac{\eta^2}{\pi}\right)^{\frac{3}{2}} e^{-\eta^2|\mathbf{r}-\mathbf{r}_j-\mathbf{nL}|^2}\,. \tag{C.2}$$

The convergence parameter η determines the decomposition of electrostatic potential into real- and Fourier-space sums, but not affects the total electrostatic energy.

C.1 Fourier space sum

The long-ranged potential due to compensating charge is expressed in Fourier space. The charge density and the electrostatic potential are given in Fourier space as

$$\begin{aligned} \rho(\mathbf{G}) \;&=\; \int_{\Omega} d\mathbf{r}\,\rho(\mathbf{r})e^{-i\mathbf{G}\cdot\mathbf{r}} = \sum_{j=1}^{N} q_j e^{-\frac{|\mathbf{G}|^2}{4\eta^2}} e^{-i\mathbf{G}\cdot\mathbf{r}_j}\,, \\ V_{\rm lr}(\mathbf{G}) \;&=\; \frac{4\pi}{|\mathbf{G}|^2}\rho(\mathbf{G})\,, \end{aligned} \tag{C.3}$$

where Ω is the volume of the cell. $\mathbf{G}=0$ is omitted in calculating electrostatic potential $V_{\rm lr}(\mathbf{G})$ since the total compensating charge is neutral. Inverse Fourier transform of the potential back to real space gives

$$\begin{aligned} V_{\rm lr}(\mathbf{r}) \;&=\; \frac{1}{\Omega}\sum_{\mathbf{G}\neq 0} V(\mathbf{G})e^{i\mathbf{G}\cdot\mathbf{r}} \\ &=\; \frac{1}{\Omega}\sum_{\mathbf{G}\neq 0}\sum_{j=1}^{N} \frac{4\pi q_j}{|\mathbf{G}|^2} e^{-\frac{|\mathbf{G}|^2}{4\eta^2}} e^{i\mathbf{G}\cdot(\mathbf{r}-\mathbf{r}_j)}\,. \end{aligned} \tag{C.4}$$

Finally, the Fourier part of Ewald sum due to the long range electrostatic potential of compensating charges is

$$\begin{aligned} E_{\rm lr} &= \frac{1}{2}\sum_{i=1}^{N} q_i V(\mathbf{r}_i) \\ &= \frac{1}{2\Omega}\sum_{\mathbf{G}\neq 0}\sum_{i,j=1}^{N} \frac{4\pi q_i q_j}{|\mathbf{G}|^2} e^{-\frac{|\mathbf{G}|^2}{4\eta^2}} e^{i\mathbf{G}\cdot(\mathbf{r}_i-\mathbf{r}_j)} \,. \end{aligned} \tag{C.5}$$

C.2 Self-interaction

The self-interaction $\frac{1}{2}q_i V_{\rm self}(\mathbf{r}_i)$ due to a point charge q_i and the compensating Gaussian function centered at $\mathbf{r}_i$ needs to be excluded from Eqn. (C.5). The potential due to single Gaussian function, $\frac{q_i}{\eta^3}(\pi)^{-\frac{3}{2}} e^{-\eta^2 r^2}$, is $V_{\rm self}(\mathbf{r}) = \frac{q_i}{r}{\rm erf}(\eta r)$ [88], where the error function ${\rm erf}(r) = \frac{2}{\sqrt{\pi}}\int_0^r du\, e^{-u^2}$. In the limit of $r \to 0$, $V_{\rm self}(r=0) = 2q_i\frac{\eta}{\sqrt{\pi}}$. Hence the self-interaction energy is $E_{\rm self} = \frac{\eta}{\sqrt{\pi}}\sum_i q_i^2$. It is a constant term.

C.3 Real space sum

The short-ranged potential is due to the point charges screened by opposite charged Gaussian functions

$$\begin{aligned} V_{\rm sr}(\mathbf{r}) &= \frac{q_i}{r} - \frac{q_i}{r}{\rm erf}(\eta\, r) \\ &= \frac{q_i}{r}\,{\rm erfc}(\eta\, r)\,, \end{aligned} \tag{C.6}$$

here, ${\rm erfc}(r) = 1 - {\rm erf}(r)$ is the complementary error function. Thus the real part of Ewald sum due to the short-ranged potential of screened charges is given by

$$E_{\rm sr} = \frac{1}{2}\sum_{i\neq j}^{N} q_i q_j \frac{{\rm erfc}(\eta\, r_{ij})}{r_{ij}} \,. \tag{C.7}$$

At the end, the total electrostatic energy is

$$
\begin{aligned}
E &= E_{\mathrm{sr}} + E_{\mathrm{lr}} - E_{\mathrm{self}} \\
&= \frac{1}{2}\sum_{i\neq j}^{N} q_i q_j \frac{\operatorname{erfc}(\eta\, r_{ij})}{r_{ij}} + \frac{1}{2\Omega}\sum_{\mathbf{G}\neq 0}\sum_{i,j=1}^{N} \frac{4\pi q_i q_j}{|\mathbf{G}|^2} e^{-\frac{|\mathbf{G}|^2}{4\eta^2}} \cos\left[\mathbf{G}\cdot(\mathbf{r}_i - \mathbf{r}_j)\right] \\
&\quad - \frac{\eta}{\sqrt{\pi}}\sum_i q_i^2 \,.
\end{aligned}
\tag{C.8}
$$

Appendix D

Compensating model potential in real and reciprocal spaces

Section 3.2.2 discussed the construction of a spherically symmetric model charge density and potential for each species of ions μ. The model charge density smoothly decays to zero as increasing r to r_c as shown in Figure 3.2. The corresponding potential is

$$V_\mu^{\text{model}}(r) = \begin{cases} \frac{1}{5r_c}\left[12 - 14u^2 + 28u^5 - 30u^6 + 9u^7\right] & : r < r_c \\ 1/r & : r > r_c \end{cases} \tag{D.1}$$

with $u = r/r_c$.

The Fourier transform of this spherically symmetric potential can be derived analytically,

$$\begin{aligned} V_\mu^{\text{model}}(\mathbf{G}) &= \int_{\text{all space}} V_\mu^{\text{model}}(\mathbf{r})\, e^{-i\mathbf{G}\cdot\mathbf{r}}\, d\mathbf{r} \\ &= \int_0^\infty r^2 dr \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi\, V_\mu^{\text{model}}(r)\, e^{-i|\mathbf{G}|r\cos\theta} \\ &= \int_0^\infty r^2 dr \int_{-1}^1 dx\, 2\pi\, V_\mu^{\text{model}}(r)\, e^{-i|\mathbf{G}|rx}\,. \end{aligned} \tag{D.2}$$

$$\begin{aligned} V_\mu^{\text{model}}(|\mathbf{G}|) =& \frac{4\pi}{|\mathbf{G}|}\int_0^\infty \sin(|\mathbf{G}|r)\, V_\mu^{\text{model}}(r)\, r\, dr \\ =& 4\pi r_c^2 \frac{-(72576 - 10080v^2)\cos v + 1008\,(72 - 4v^2 + v(-42 + v^2)\sin v)}{v^{10}}\,, \end{aligned} \tag{D.3}$$

where $v = |\mathbf{G}|r_c$. To avoid the numerical truncation error, we expand $V_\mu^{\text{model}}(|\mathbf{G}|)$ as a polynomial for a small $|\mathbf{G}|$ (e.x. $|\mathbf{G}| < 0.5$), so that the final estimation of $V_\mu^{\text{model}}(|\mathbf{G}|)$ can

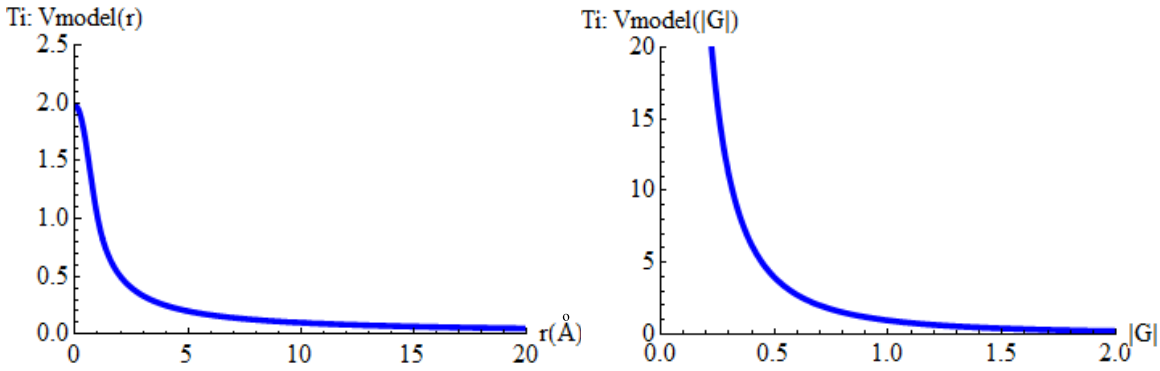

Figure D.1: Compensating model potential in real and reciprocal spaces for PAW Ti with a cutoff radius 1.22 Å.

be in double precision,

$$V_{\mu}^{\text{model}}(|\mathbf{G}|) = 4\pi r_c^2 \left(\frac{1}{v^2} - \frac{7}{150} + \frac{v^2}{1100} - \frac{v^4}{98280} + \frac{v^6}{13305600} - \frac{v^8}{2520460800} + O[v]^{10} \right) . \tag{D.4}$$

References

[1] M. Haruta, T. Kobayashi, H. Sano, and N. Yamada, Novel gold catalysts for the oxidation of carbon monoxide at a temperature far below 0 °C, Chem. Lett. **2**, 405 (1987).

[2] M. Haruta, N. Yamada, T. Kobayashi, and S. Iijima, Gold catalysts prepared by coprecipitation for low-temperature oxidation of hydrogen and of carbon monoxide, J. Catal. **115**, 301 (1989).

[3] M. Valden, X. Lai, and D. W. Goodman, Onset of catalytic activity of gold clusters on titania with the appearance of nonmetallic properties, Science **281**, 1647 (1998).

[4] F. Boccuzzi, A. Chiorino, M. Manzoli, D. Andreeva, and T. Tabakova, FTIR study of the low-temperature water-gas shift reaction on Au/Fe_2O_3 and Au/TiO_2 catalysts, J. Catal. **188**(1), 176 (1999).

[5] T. Hayashi, K. Tanaka, and M. Haruta, Selective vapor-phase epoxidation of propylene over Au/TiO_2 catalysts in the presence of oxygen and hydrogen, J. Catal. **178**, 566 (1998).

[6] U. Diebold, The surface science of titanium dioxides, Surf. Sci. Rep. **48**, 53 (2003).

[7] G. C. Bond and D. T. Thompson, Gold-catalysed oxidation of carbon monoxide, Gold Bull. **33**(2), 41 (2000).

[8] J.-D. Grunwaldt, M. Maciejewski, O. Becker, P. Fabrizioli, and A. Baiker, Comparative Study of Au/TiO_2 and Au/ZrO_2 Catalysts for Low-Temperature CO Oxidation, J. Catal. **186**, 458 (1999).

[9] M. Haruta, Gold as a novel catalyst in the 21^{st} century: preparation, working mechanism and applications, Gold Bull. **37**(1-2), 27 (2004).

[10] F. Cosandey and T. E. Madey, Growth, morphology, interfacial effects and catalytic properties of Au on TiO_2, Surf. Rev. Lett. **8**, 73 (2001).

[11] S. Sivaramakrishnan, B. J. Pierce, M. E. Scarpelli, M. Yu, J. Wen, B. Jiang, D. R. Trinkle, and J. M. Zuo, Enhanced adhension of gold nanocrystals on rutile surfaces via interfacial reconstructions, submitted .

[12] T. Akita, K. Tanaka, M. Kohyama, and M. Haruta, HAADF-STEM observation of Au nanoparticles on TiO_2, Surf. Interface Anal. **40**, 1760 (2008).

[13] W. L. Winterbottom, Equilibrium shape of a small particle in contact with a foreign substrate, Acta Metall. **15**(2), 303 (1967).

[14] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. A **140**, 1133 (1965).

[15] N. Chetty and R. M. Martin, First-principles energy density and its applications to selected polar surfaces, Phys. Rev. B **45**, 6074 (1992).

[16] N. Chetty, First principle energy density and its applications to selected polar surfaces and interfaces, Ph.D. Thesis, University of Illinois at Urbana-Champaign (1990).

[17] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press: Oxford, 1990.

[18] J. I. Rodríguez, P. W. Ayers, A. W. Götz, and F. L. Castillo-Alvarado, Virial theorem in the Kohn-Sham density-functional theory formalism: Accurate calculation of the atomic quantum theory of atoms in molecules energies, J. Chem. Phys. **131**(2), 021101 (2009).

[19] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, Phys. Rev. B **136**, 864 (1964).

[20] J. P. Perdew and A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, Phys. Rev. B **23**(10), 5048 (1981).

[21] J. P. Perdew and Y. Wang, Accurate and simple analytic representation of the electron-gas correlation energy, Phys. Rev. B **45**(23), 13244 (1992).

[22] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. **77**(18), 3865 (1996).

[23] D. R. Hamann, M. Schlüter, and C. Chiang, Norm-Conserving Pseudopotentials, Phys. Rev. Lett. **43**(20), 1494 (1979).

[24] D. Vanderbilt, Soft self-consistent pseudopotentials in a generalized eigenvalue formalism, Phys. Rev. B **41**(11), 7892 (1990).

[25] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**(24), 17953 (1994).

[26] D. M. Ceperley and B. J. Alder, Ground state of the electron gas by a stochastic method, Phys. Rev. Lett. **45**(7), 566 (1980).

[27] S. H. Vosko, L. Wilk, and M. Nusair, Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis, Can. J. Phys. **58**(8), 1200 (1980).

[28] O. Gunnarsson and B. I. Lundqvist, Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism, Phys. Rev. B **13**(10), 4274 (1976).

[29] R. O. Jones and O. Gunnarsson, The density functional formalism, its applications and prospects, Rev. Mod. Phys. **61**(3), 689 (1989).

[30] W. E. Pickett, Pseudopotential methods in condensed matter applications, Comput. Phys. Rep. **9**, 115 (1989).

[31] K. Laasonen, A. Pasquarello, R. Car, C. Lee, and D. Vanderbilt, Car-Parrinello molecular dynamics with Vanderbilt ultrasoft pseudopotentials, Phys. Rev. B **47**(16), 10142 (1993).

[32] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys. Rev. B **59**(3), 1758 (1999).

[33] N. Troullier and J. L. Martins, Efficient psedopotentials for plane-wave calculations, Phys. Rev. B **43**(3), 1993 (1991).

[34] L. Kleinman and D. M. Bylander, Efficacious form for model pseudopotentials, Phys. Rev. Lett. **48**(20), 1425 (1982).

[35] X. Gonze, P. Käckell, and M. Scheffler, Ghost states for separable, norm-conserving, *ab initio* pseudopotentials, Phys. Rev. B **41**(17), 12264 (1990).

[36] S. G. Louie, S. Froyen, and M. L. Cohen, Nonlinear ionic pseudopotentials in spin-density-functional calculations, Phys. Rev. B **26**(4), 1738 (1982).

[37] P. P. Ewald, Die Berechnung optischer und elektrostatischer Gitterpotentiale, Ann. Phys. **369**, 253 (1921).

[38] G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**(16), 11169 (1996).

[39] K. Rapcewicz, B. Chen, B. Yakobson, and J. Bernholc, Consistent methodology for calculating surface and interface energies, Phys. Rev. B **57**(12), 7281 (1998).

[40] J. C. Slater, Wave functions in a periodic potential, Phys. Rev. **51**(10), 846 (1937).

[41] R. M. Martin, *Electronic structure: Basic theory and practical methods*, Cambridge University Press, 2004.

[42] F. W. Biegler-König, T. T. Nguyen-Dang, Y. Tal, R. F. W. Bader, and A. J. Duke, Calculation of the average properties of atoms in molecules, J. Phys. B: At. Mol. Phys. **14**(16), 2739 (1981).

[43] P. L. A. Popelier, MORPHY, a program for an automated "atoms in molecules" analysis, Comput. Phys. Commun. **93**, 212 (1996).

[44] N. O. J. Malcolm and P. L. A. Popelier, An algorithm to delineate and integrate topological basins in a three-dimensional quantum mechanical density function, J. Comput. Chem. **24**, 1276 (2003).

[45] B. P. Uberuaga, E. R. Batista, and H. Jónsson, Elastic sheet method for identifying atoms in molecules, J. Chem. Phys. **111**, 10664 (1999).

[46] G. Henkelman, A. Arnaldsson, and H. Jónsson, A fast and robust algorithm for Bader decomposition of charge density, Comput. Mater. Sci. **36**(3), 354 (2006).

[47] W. Tang, E. Sanville, and G. Henkelman, A grid-based Bader analysis algorithm without lattice bias, J. Phys. Condens. Matter **21**(8), 084204 (2009).

[48] A. Otero-de-la-Roza and V. Luaña, A fast and accurate algorithm for QTAIM integration in solids, J. Comput. Chem. (2010), (in press).

[49] G. Voronoi, Nouvelles applications des paramètres continus à la théorie des formes quadratiques, J. reine angew. Math. **133**, 97–178 (1907).

[50] N. Chetty and R. M. Martin, GaAs (111) and $(\bar{1}\bar{1}\bar{1})$ surfaces and the GaAs/AlAs (111) heterojunction studied using a local energy density, Phys. Rev. B **45**, 6089 (1992).

[51] R. Ramprasad, First-principles energy and stress fields in defected materials, J. Phys.: Condens. Matter **14**, 5497 (2002).

[52] H. J. Monkhorst and J. D. Pack, Special points for Brillouin-zone integrations, Phys. Rev. B **13**(12), 5188 (1976).

[53] N. Moll, A. Kley, E. Pehlke, and M. Scheffler, GaAs equilibrium crystal shape from first principles, Phys. Rev. B **54**(12), 8844 (1996).

[54] G. Qian, R. M. Martin, and D. J. Chadi, First-principles calculations of atomic and electronic structure of the GaAs(110) surface, Phys. Rev. B **37**(3), 1303 (1988).

[55] R. Choudhury, D. R. Bowler, and M. J. Gillan, Atomic structure of misfit dislocations at InAs/GaAs(110), J. Phys.: Condens. Matter **20**(23), 235227 (2008).

[56] C. Messmer and J. C. Bilello, The surface energy of Si, GaAs, and GaP, J. Appl. Phys. **52**(7), 4623 (1981).

[57] A. F. Wright, Density-funcitonal-theory calculations for the silicon vacancy, Phys. Rev. B **74**(16), 165116 (2006).

[58] M. J. Puska, S. Pöykkö, M. Pesola, and R. M. Nieminen, Convergence of supercell calculations for point defects in semiconductors: Vacancy in silicon, Phys. Rev. B **58**(3), 1318 (1998).

[59] S. Dannefaer, P. Mascher, and D. Kerr, Monovacancy Formation Enthalpy in Silicon, Phys. Rev. Lett. **56**(20), 2195 (1986).

[60] E. O. Kane, Charge disturbance around an impurity in silicon, Phys. Rev. B **31**(8), 5199 (1985).

[61] H. L. McMurry, A. W. Solbrig Jr., and J. K. Boyter, The use of valence force potentials in calculating crystal vibrations, J. Phys. Chem. Solids **28**(12), 2359 (1967).

[62] S. A. Centoni, B. Sadigh, G. H. Gilmer, T. J. Lenosky, T. D. de la Rubia, and C. B. Musgrave, First-principles calculation of intrinsic defect formation volumes in silicon, Phys. Rev. B **72**(19), 195206 (2005).

[63] R. J. Needs, First-principles calculations of self-interstitial defect structures and diffusion paths in silicon, J. Phys.: Condens. Matter **11**(50), 10437 (1999).

[64] A. Ural, P. B. Griffin, and J. D. Plummer, Self-diffusion in silicon: Similarity between the properties of native point defects, Phys. Rev. Lett. **83**(17), 3454 (1999).

[65] R. G. Hennig, D. R. Trinkle, J. Bouchet, S. G. Srinivasan, R. C. Albers, and J. W. Wilkins, Impurities block the α to ω martensitic transformation in titanium, Nat. Mater. **4**(2), 129 (2005).

[66] R. M. Wood, The lattice constants of high purity alpha titanium, Proc. Phys. Soc. **80**, 783 (1962).

[67] R. J. Weiss, Compton profile of chromium: Fermi momenta of transition metals, Philos. Mag. **27**(6), 1461 (1973).

[68] O. Jepsen, Electronic structure and magnetic breakdown in titanium, Phys. Rev. B **12**(8), 2988 (1975).

[69] Z. Yang, R. Wu, and D. W. Goodman, Structural and electronic properties of Au on TiO_2(110), Phys. Rev. B **61**(20), 14066 (2000).

[70] Y. Wang and G. S. Hwang, Adsorption of Au atoms on stoichiometric and reduced TiO_2(110) rutile surfaces: a first principles study, Surf. Sci. **542**(1), 72 (2003).

[71] V. Amrendra, M. Greg, and M. Horia, Adsorption of gold on stoichiometric and reduced rutile TiO_2(110) surfaces, J. Chem. Phys. **118**, 6536 (2003).

[72] E. Wahlström, N. Lopez, R. Schaub, P. Thostrup, A. Rønnau, C. Africh, E. Lægsgaard, J. K. Nørskov, and F. Besenbacher, Bonding of gold nanoclusters to oxygen vacancies on rutile TiO_2(110), Phys. Rev. Lett. **90**(2), 026101 (2003).

[73] K. Okazaki, Y. Morikawa, S. Tanaka, K. Tanaka, and M. Kohyama, Electronic structures of Au on TiO_2(110) by first-principles calculations, Phys. Rev. B **69**(23), 235404 (2004).

[74] N. Lopez, J. K. Nørskov, T. V. W. Janssens, A. Carlsson, A. Puig-Molina, B. S. Clausen, and J.-D. Grunwaldt, The adhesion and shape of nanosized Au particles in a Au/TiO_2 catalyst, J. Catal. **225**, 86 (2004).

[75] T. Pabisiak and A. Kiejna, First-principles study of Au nanostructures on rutile TiO_2(110), Phys. Rev. B **79**(8), 085411 (2009).

[76] D. Pillay and G. S. Hwang, Growth and structure of small gold particles on rutile TiO_2(110), Phys. Rev. B **72**(20), 205422 (2005).

[77] H. Shi, M. Kohyama, S. Tanaka, and S. Takeda, Structure and stability of Au rods on TiO_2(110) surfaces by first-principles calculations, Phys. Rev. B **80**(15), 155413 (2009).

[78] N. Shibata, A. Goto, K. Matsunaga, T. Mizoguchi, S. D. Findlay, T. Yamamoto, and Y. Ikuhara, Interface structures of gold nanoparticles on TiO_2(110), Phys. Rev. Lett. **102**(13), 136105 (2009).

[79] C. L. Pang, S. A. Haycock, H. Raza, P. W. Murray, G. Thornton, O. Gülseren, R. James, and D. W. Bullett, Added row model of TiO_2(110) 1×2, Phys. Rev. B **58**(3), 1586 (1998).

[80] A. Kiejna, T. Pabisiak, and S. W. Gao, The energetics and structure of rutile TiO_2(110), J. Phys.: Condens. Matter **18**, 4207 (2006).

[81] M. Ramamoorthy, D. Vanderbilt, and R. D. King-Smith, First-principles calculations of the energetics of stoichiometric TiO_2 surfaces, Phys. Rev. B **49**(23), 16721 (1994).

[82] G. Charlton, P. B. Howes, C. L. Nicklin, P. Steadman, J. S. G. Taylor, C. A. Muryn, S. P. Harte, J. Mercer, R. McGrath, D. Norman, T. S. Turner, and G. Thornton, Relaxation of TiO_2(110)-(1×1) using surface x-ray diffraction, Phys. Rev. Lett. **78**(3), 495 (1997).

[83] R. Lindsay, A. Wander, A. Ernst, B. Montanari, G. Thornton, and N. M. Harrison, Revisiting the surface structure of TiO_2(110): A quantitative low-energy electron diffraction study, Phys. Rev. Lett. **94**(24), 246102 (2005).

[84] M. Methfessel and A. T. Paxton, High-precision sampling for Brillouin-zone integration in metals, Phys. Rev. B **40**(6), 3616 (1989).

[85] V. Zólyomi, L. Vitos, S. K. Kwon, and J. Kollár, Surface relaxation and stress for 5d transition metals, J. Phys.: Condens. Matter **21**, 095007 (2009).

[86] J. C. W. Swart, P. v. Helden, and E. v. Steen, Surface energy estimation of catalytically relevant fcc transition metals using DFT calculations on nanorods, J. Phys. Chem. C **111**(13), 4998 (2007).

[87] A. D. Becke and K. E. Edgecombe, A simple measure of electron localization in atomic and molecular systems, J. Chem. Phys. **92**(9), 5397 (1990).

[88] D. Frenkel and B. Smit, *Understanding molecular simulation: From algorithms to applications*, Academic Press, 2002.